

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTA
TGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAA**ATG**TGG
TGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTC
ATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGT
GACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTT
TTATGCATTGCTACCATTTATGTTGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAAC
GTTATCATCAAATTAACAAGGCTGGCCTTGTAAGTACTGAGTTGTTTAGGACTTTCT
ATTGTGGCAAACCTCCAGAAAAACAACCCTTTTTGCTGCACATGTAAGTGGAGCTGTGCTTACC
TTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAA
ATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCA
CTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAG
AAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAA
TGGTCTATGTCATTTTCCTTCTTTGGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATT
TCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTGCACCTTGCCCTATT
AACAAATGAACGAACACGGCTACTTTCCAGAGATATTT**TGAT**GAAGGATAAAATATTTCTGTAA
TGATTATGATTCTCAGGGATTGGGGAAAGGTTACAGAAAGTTGCTTATTCTTCTCTGAAATTT
TCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATAATCAGGAAACATGAAAG
AAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTATTAAAAACACCTATGCCT
ATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIIFSITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIA
AVLCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAV
LTFGMGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDL
EQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPC
PINNERTRLLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGGC
AGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGC
AGGAGCCTTCCTTACACTTCGCC**ATG**AGTTTCCTCATCGACTCCAGCATCATGATTACCTCCC
AGATACTATTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGA
TACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGCACCATGTTTG
AGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAA
TGAACCTGTGTGAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTACATTGGCTATTTTA
TTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTTCCTGTCTCTTATGGCTGA
CCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCCCATTTCTCAGCCCAAACATGGGA
TCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAGTGACTCTCATGGCTCTTC
TTTCTGGATTTGGTGCTGTCAACTGCCCATACACTTACATGTCTTACTTCCTCAGGAATGTGA
CTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGATCATAAGCA
AAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACAAAC
CATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGGAAGTGAAAATCTTA
CTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAG
CTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTTA
ATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATA
TTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCACAAGAGGCATTGAGATCACTGTGAATT
ATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCCAACACATTTCTTCATTCTTGTGGAA
TAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTACCAAGTTCTTTTATGCCATCT
CTAGCAGTAAGTCCTCCAATGTCATTGTCTGCTATTAGCACAGATAATGGGCATGTACTTTG
TCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAATACCGCACCATAATCACTGAAG
TCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTGATGTGATCTTCCTGGTCAGCG
CTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAACAGGCACCAGAGAAGCAAATGGCAC
CT**TGA**ACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAATTTAGATATAAGAGG
GGGGAAAAATGGAACCAGGGCCTGACATTTTATAAACAAACAAAATGCTATGGTAGCATTTTT
CACCTTCATAGCATACTCCTTCCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAGA
ACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGG
CTGGTGTAGAGGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGC
AAGACATGTCTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTAAGGTTACATGGAAA
AGGTTATAGCTTTGCCTTGAGATTGACTCATTAATAATCAGAGACTGTAACAAAAAAAAAAAA
AAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCACT
TGTTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFFGWLFMRQLFKDYEIRQYVVQVIFSVTFASFCTMFELIIFEIL
GVLNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWK
LGDPPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILAL
ERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVD
ALEELSRQLFLETADLYATKERIEYSKTFKGKYFNLGYFFSIYCVWKIFMATINIVFDRVKG
TDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNV
IVLLLAQIMGYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFL
YLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTT
 CCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTG
 AGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACG**ATG**GCCAGGTGCTTCAGCCTGGTGTGCTT
 CTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCC
 ATCCAGGTGTCATGCAGAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCAGCAGCTG
 AATTTACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAA
 GTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGGATGGATTCT
 GTGGTCATCTCTAGGATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCCTGATT
 TGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACT
 AACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATTCAACACTCAAACCTGCAACA
 CAAACAACAGAATTTATTGTGAGTGACAGTACCTACTCGGTGGCATCCCCTTACTCTACAATA
 CCTGCCCCCTACTACTCTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTG
 ATTTGTGTCACAGAAGTTTTTATGGAACTAGCACCATGTCTACAGAACTGAACCATTTGTT
 GAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTA
 GTGCTTGCTCTCCTCTTCTTTGGTGCTGCAGCTGGTCTTGGATTTTGCTATGTCAAAGGTAT
 GTGAAGGCCTTCCCTTTTACAAACAAGAATCAGCAGAAGGAAATGATCGAAACCAAAGTAGTA
 AAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGAAAACTGATAAAAACCCA
 GAAGAGTCCAAGAGTCCAAGCAAACTACCGTGCGATGCCTGGAAGCTGAAGTT**TAG**ATGAGA
 CAGAAATGAGGAGACACACCTGAGGCTGGTTTTCTTTTCATGCTCCTTACCCTGCCCCAGCTGGG
 GAAATCAAAGGGCCAAAGAACCAAAGAAGAAAGTCCACCCTTGGTTCTCTAACTGGAATCAGC
 TCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGCCCTTCTCCTTATTGTAACCCT
 GTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCTTTCTAGCCTGGCTATGTCCTA
 ATAATATCCCCTGGGAGAAAGGAGTTTTTGCAAAGTGCAAGGACCTAAACATCTCATCAGTA
 TCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTGAAAGCCAAGGAGTCACT
 GAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGACCCCTTCTTCAGCTCTGAAAGAGAAA
 CACGTATCCCACCTGACATGTCCTTCTGAGCCCGGTAAGAGCAAAAGAATGGCAGAAAAGTTT
 AGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAAGCTAAAATAAA
 GAAATAGAACAAAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAACACAGACAGG
 GTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACACTTACTT
 TTTCTGGTCTCTACCACTGCTGATATTTTCTCTAGGAAATATACTTTTACAAGTAACAAAAAT
 AAAAATCTTATAAATTTCTATTTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATTACT
 CAGTAATTTGTTTAAAAAGTAATAAAATTCAACAAACATTTGCTGAATAGCTACTATATGTCA
 AGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATTGCACATAGTA
 GAACGCTATCTGGGAAGCTATTTTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACT
 AATTTTTATTTTTGCTGAGACTAATCTTATTCATTTTTCTCTAATATGGCAACCATTATAACCT
 TAATTTATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAGCACATTTTAA
 AAGTGCCATTAACAAATGTATCACTAGCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGC
 AGAAATATTTGTGACAAAAAATTAAAGCATTTAGAAAATT

FIGURE 6

MARCFSLVLLLLTSIWTTLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRL
LGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAA
YCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTTPPAPA
STSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAA
GLGFCYVKRYVKAFFFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPESKSPSKTTV
RCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGCGAGCCCG
GCGGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCGGG
GCGGCGGCTGCGGGCGCAGAGCGGAG**ATG**CAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGC
TGGCGGCGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGC
CCGGCCCCGGCTCTCAGCTACCCGCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTG
AGGAACTGATGGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAG
AAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCCAGCTATCACAATG
AGACCAACACAGACACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGA
TAACCAACAACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACG
AGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGCATGTACTGCC
AGTTTGCCAGCTTCCAGTACACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGG
ACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAATGGCCACCAGGG
GCAGCAATGGGACCATCTGTGACAACCAGAGGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCC
AGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACC
CCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGAT
GCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTGGTGTATGTGTGCAAGC
CGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCCAGAGAGGTCCCCGATG
AGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGGAGGACCTGGAGAGGAGCC
TGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCCGCTGCACTGCTGGGAGGGGAAG
AGATT**TAG**ATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAAATAGCTAATTTATTTCCCCA
GGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCTACATCTTCTTCCAGTAAGTTTCC
CCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGCTCCCCAGGCTGTTCTCCAGGC
TTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAACTGCAGGAGCAGTTTGCCACCCCT
GTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAGACAGCCGTTTGTCTACATGGCTT
TGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGGAGTCTCCCTCTGATTGGTTTTGGGG
AAATGTGGAGAAGAGTGGCCTGCTTTGCAAACATCAACCTGGCAAAAATGCAACAAATGAATT
TTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGCAGATGAAATGT
TCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTTGCTCAGCTCCTACCTCTGT
GCCAGGGCAGCATTTTCATATCCAAGATCAATTCCTCTCTCAGCACAGCCTGGGGAGGGGGT
CATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCCCAAGT
CACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCTCTC
CACTACCCACACACCAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGATT
TTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACCTAATTCTCACATCCCTCTAAAAGTAA
ACTACTGTTAGGAACAGCAGTGTTCTCACAGTGTTGGGGCAGCCGTCCTTCTAATGAAGACAAT
GATATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACTTTGAAAGGTATATGACTGAGCG
TAGCATACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGA
AATTTGCAAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACC
GAGCAGGGCTGTGTGAAACATGGTTGTAATATGCGACTGCGAACAACCTGAACCTCTACGCCACTC
CACAAATGATGTTTTAGGTGTCATGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAA
GTTTAAAGTTGCACATGATTGTATAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACAT
AAGTTGCATTTAGAAATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 8

MQRLGATLLCLLLAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQH
KLRSVEEMEAEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQM
VFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRM LCTRDSECCGDQL
CVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRLDLI
TWELEPDGALDRPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFME
EVRQELEDLERSLTEEMALGEPAAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCC
ACGGCCACCTTGTGAACTCCTCGTGCCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCC
AAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTTGGGGCTCTTCTGG
ACCCTTAACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTTCTAC
TGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACA
CTCCGTTACCACACTGGGTCATTGGCATTGAGAGCCCTCATCCTGACCCTTGTGCAGATAGCC
CGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCCGCTGC
ATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGC
AATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTC
ATGCTACTCATGCGAAACATTGTGAGGGTGGTCTGCTGACAAAGTCACAGACCTGCTGCTG
TTCTTTGGGAAGCTGCTGGTGGTGGAGGCGTGGGGGCTCTGCTCCTTCTTTTTTTCTCCGGT
CGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATC
ATGACCTCCATCCTGGGGGCTATGTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGT
GTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGG
CCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAACGAGGCGCCCCGGAC
AACAAGAAGAGGAAGAAG**TGA**CAGCTCCGGCCCTGATCCAGGACTGCACCCACCCCCACCGT
CCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGTAAAAAAGGTTTTAGGC
CAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTGAGGCGGGCGGATCACCTG
AGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCCGTCTCTATTAAAAATACAA
AAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTACTCGGGAGGCTGAGGCAGGAG
AATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAACC
TGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAAAGATTTTATTAAAGATATTT
TGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFW
TLNWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIA
RVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAF
MLLMRNIVRVVLDKVTDLLLFFGKLLVVGGVGVLSSFFFSGRIPGLGKDFKSPHLNYYWLP
MTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPD
NKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAG
 CCTGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCT
 GCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCCTGG
 GGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGAGAGTCAGCTCTACAAGCTGCCCT
 GGGTGCTGAGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGGCTCCC
 TGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTCTTT
 TCACCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAGAATGGGT
 TTTGGTTCTTTAAGTTCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACATCCCTGACG
 GCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCGTGGGCTCCTTCCTCTTCATCCTCATCC
 AGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGCAAGGCCGAGG
 AGTGCGATTCCCGTGCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTTCTACTTGCTGT
 CGATCGCGGCCGTGGCGCTGATGTTTCATGTACTACACTGAGCCAGCGGCTGCCACGAGGGCA
 AGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGCTGTCTGCCCCA
 AGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTTCATCACCTCTACACCA
 TGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTTGCCAA
 CCCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATG
 CCCCAGCATTGTGGGCCTCATCATCTTCCTCCTGTGCACCCTCTTCATCAGTCTGCGCTCCT
 CAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCA
 CACAGCAGCAGCAGCAGGCTGGCAGCCTGTGAGGGCCGGGCCTTTGACAACGAGCAGGACG
 GCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGTGGCCTCACTGCACGTCTATGA
 TGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAAGATGATCAGCACGTGGACCGCCG
 TGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTCTACCTGTGGACCCTGGTAGCCC
 CACTCCTCCTGCGCAACCGCGACTTCAGCTGAGGCAGCCTCACAGCCTGCCATCTGGTGCCTC
 CTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTCCCCACACCAATCAGCC
 AGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTGCCCCCTGAGCCGGGCCTTCTAGTCGT
 AGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCACACCCAC
 ACGGTGGAGCTGCCTCTTCTTCCCCTCCTCCCTGTTGCCATACTCAGCATCTCGGATGAAA
 GGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACCTCCCACC
 ACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCAGGGGACCCTGCC
 CCCTTCTGGACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTCTAATAAACAAGCCAGTGCG
 TGTAACAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLEFLGVLVSIIMLSPGVES
QLYKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPR
AAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQR
WLGKAEECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVS
IAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGY
ETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRA
FDNEQDGVTYSSFFHFCLVLASLHVMMTLTNWYKPGETRMISTWTAVVVKICASWAGLLLY
LWTLVAPLLLRNRDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGT
GAGGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGG**ATGA**ACCACCTGCCAGAAGACATGGAG
AACGCTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCC
ACACAACCTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTC
AGGAGGACTTTCTGTTTGTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATA
GAGTTAAATGTGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTAC
TATTCTTCATATTTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCA
TATGCTGTGTGCAGACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCC
TTTTTACTAGCAAAAGTGATCCTTTTGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTG
CCCATCATTTTCATTCATCCTTGCCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCT
CAAGAAGCAGAAGAAGAAAACAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTT
ATACCTGGTGGTCTTTCTGATGGTCAGTTTTATTCCCCTCCTGAATCCGAAGCAGGATCTGAA
GAAGCTGAAGAAAAACAGGACAGTGAGAAACCACTTTTAGAACTAT**TGA**GTA CTACTTTTGTTA
AATGTGAAAAACCCTCACAGAAAGTCATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTG
TCGACAGTAAAGTTGAAATGGTGACGTCCACTGCTGGCTTTATTGAACAGCTAATAAAGATTT
ATTTATTGTAATACCTCACAAACGTTGTACCATATCCATGCACATTTAGTTGCCTGCCTGTGG
CTGGTAAGGTAATGTCATGATTCATCCTCTCTTCAGTGAGACTGAGCCTGATGTGTTAACAAA
TAGGTGAAGAAAGTCTTGTGCTGTATTCCTAATCAAAGACTTAATATATTGAAGTAACACTT
TTTTAGTAAGCAAGATACCTTTTTTATTTCAATTCACAGAATGGAATTTTTTTGTTTCATGTCT
CAGATTTATTTTGTATTTCTTTTTTAACACTCTACATTTCCCTTGTTTTTTAACTCATGCACA
TGTGCTCTTTGTACAGTTTTAAAAAGTGTAATAAAATCTGACATGTCAATGTGGCTAGTTTTA
TTTTTCTTGTTTTGCATTATGTGTATGGCCTGAAGTGTGGACTTGCAAAGGGGAAGAAAGG
AATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTTTTATCATGAAATCATGTTT
TTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGCACAAAATGACTTAAACC
ATTCATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAAATGAATAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDL
LFVTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAI
ALTTAVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQEAEEENRLLIV
QDASERAALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCCGACCCGCCAGGAAAGACTGA
 GGCCGCGGCCTGCCCCGCGGCTCCCTGCGCCGCCGCCCTCCCGGGACAGAAG**ATGT**GCT
 CCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCTGGGGTGCAGGGCT
 GCCATCCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCA
 CGGTGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCA
 TGCTCGACGCAGGCAGCTTTGCCGGCCTGCCGGGCTGCAGCTCCTGGACCTGTACAGAACC
 AGATCGCCAGCCTGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGA
 CGGCCAACAGGCTGCATGAAATCACCATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCC
 TCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCC
 TCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCGCTGCGCCTGCCCCGCC
 TGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTG
 CCAACGTGGAGGCGCTGCGGCTGGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTCA
 GCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTG
 TGATCCGAGGCCTCCGGGGCCTGACGCGCCTGCGGCTGGCCGGCAACACCCGCATTGCCACGC
 TGCGGGCCCGAGGACCTGGCCGGCCTGGCTGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCC
 TGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCCGCCTGCGGCTGCTGGCAGCTGCCC
 GCAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCCCTGGGTGCGCGAGAGCCACG
 TCACACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCAAGAACGCTGGCCGGCTGC
 TCCTGGAGCTTGACTACGCCGACTTTGGCTGCCACGACCACCACACAGCCACAGTGCCCA
 CCACGAGGCCCGTGGTGCGGGAGCCACAGCCTTGTCTTAGCTTGGCTCCTACCTGGCTTA
 GCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTCCACTGCCCCACCGACTGTAGGGC
 CTGTCCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCAATGGGGGCACATGCCACCTGG
 GGACACGGCACACCTGGCGTGCTTGTGCCCCGAAGGCTTCACGGGCCTGTACTGTGAGAGCC
 AGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCACGCCGAGGCCACCACGGTCCCTGA
 CCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGCGGCTACCTCCAGG
 GGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCGGGCCCTGATAAGC
 GGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACACGGTCACCCAGCTGCGGCCCA
 ACGCCACTTACTCCGTCTGTGTATGCCTTTGGGGCCCGGGCGGGTGCCGGAGGGCGAGGAGG
 CCTGCGGGGAGGGCCATACACCCCCAGCCGTCCACTCCAACCACGCCCCAGTCACCCAGCCC
 GCGAGGGCAACCTGCCGCTCCTCATTGCGCCCCGCCCTGGCCGCGGTGCTCCTGGCCGCGCTGG
 CTGCGGTGGGGGACGCTACTGTGTGCGGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGGACA
 AAGGGCAGGTGGGGCCAGGGGCTGGGCCCTGGAAGTGGAGGGAGTGAAGGTCCCCTTGAGC
 CAGGCCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGC
 CACTCATGGGCTTCCAGGGCCTGGCCTCCAGTCAACCCCTCCACGCAAAGCCCTACAT**CTAAG**
 CCAGAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTG
 CTGCCACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGAC
 CACAGCTGGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCCAGC
 TGACGAGCCCTAACGTCCCCAGAACCAGAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGC
 AACGTGCAGTCCCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTCTGC
 TGGGCTCTCCCACTCCAGGCGGACCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAG
 GGAGAGCGGGTAGGCGGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAA
 ACTGGAAGGAAGATGCTTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAG
 ATCCTTTCCCATTTATTCTGGGAAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTG
 TAAGACAAACGATGATATGAAGGCCTTTTGTAAGAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLLPLLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPDPDTVGLYVFENG
 ITMLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRL
 ERLYLGNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGIL
 DTANVEALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRI
 AQLRPEDLAGLAALQELDVSNSLSLQALPGDLSGLFPRLRLAAARNPFNCVCPLSWFGPWVRE
 SHVTLASPEETRCHFPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPT
 WLSPTAPATEAPSPSTAPPTVGPVPQPDCCPSTCLNGGTCHLGRHHLACLCPEGFTGLYC
 ESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPSTSLRVGLQRYLQGSSVQLRSLRLTYRNLSGP
 DKRLVTLRLPASLAEYTVTQLRPNATYSVCMPLGPGRVPEGEEACGEAHTPPAVHSNHAPVT
 QAREGNLPLLIAPALAAVLLAALAAVGAAYCVRGRAMAAAAQDKGQVGPAGPLELEGVKVP
 LEPGPKATEGGGEALPSGSECEVPLMGFPGLQSPHAKPYI

Important features:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
 594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATG
 CGGGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCC
 TCGGATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAG
 TCAGTAAAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTGAGAA
 GAATCTGAATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAA
 AGTGTACAGAAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAG
 CCAAAGAAAGTACGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGC
 CACTTCCCTTTTCTTTTCTTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGGAAGAT
 GGCAGACTGTGGTGTGCTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAA
 ACTGAAGAAGAGGCTGCTAAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAAGTGA
 ATGAAAATCCTTAATGGAAGCAATAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAA
 AAGGCAGCAAGCATGAACCATAACCAAAGCCCTGGAGAGAGTGTATATGCTCTTTTATTTGGT
 GATTACTTGCCACAGAATATCCAGGCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGC
 TCTCCCAAGGGACAGACTGCTCTTGGCTTTCTGTATGCCTCTGGACTTGGTGTAAATTCAAGT
 CAGGCAAAGGCTCTTGTATATTATACATTTGGAGCTCTTGGGGCAATCTAATAGCCACATG
 GTTTTGGTAAGTAGACTTTAGTGAAGGCTAATAATATTAACATCAGAAGAATTTGTGGTTTA
 TAGCGGCCACAACCTTTTTCAGCTTTCATGATCCAGATTTGCTTGTATTAAGACCAAATATTCA
 GTTGAACCTCCTTCAAATTCTTGTAAATGGATATAACACATGGAATCTACATGTAAATGAAAG
 TTGGTGGAGTCCACAATTTTTCTTTAAAATGATTAGTTTGGCTGATTGCCCCATAAAAGAGAG
 ATCTGATAAATGGCTCTTTTTTAAATTTTCTCTGAGTTGGAATTGTCAGAATCATTTTTTACAT
 TAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCAAATTTTGTAAATGGTGGCTATA
 GAAAAACAACATGAAATATTATACAATTTTTGCAACAATGCCCTAAGAATTGTTAAAATTCA
 TGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTGTTTTTTACTTTTCATGATT
 GGCTGTCTTCCATTTATTCTGGTCATTTATTGCTAGTGACACTGTGCCTGCTTCCAGTAGTC
 TCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGGAAGATTAACATTTT
 TTAATAAAATTATGTCTAAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLDKTTLTSDSVKDHTTAGRVVAGQIFLDS
EESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEP
CHFPFLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAARRQMGEAEMMYQT
GMKILNGSNKKSQKREAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMFEEKLTEE
GSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLI AHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAGCCCATTTCTGCAGTGGAAATTTTCATGAACTAGCAAGAGGACACCATCTTC
 TTGTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAA**ATG**CTCTTTTGGGTGC
 TAGGCCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAGACA
 TCACTGATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAA
 CTTTTGATAAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTT
 TAAAGGCAGAAACCTCAGAGAGACTTCGTA CTGTGCTTCTGGATGTGACCGACCCAGAGAATG
 TCAAGAGGACTGCCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCA
 ATAATGCTGGTGTTCCTGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAG
 AACCTATTGAAGTGAACCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCA
 AGAAAGCTCAAGGGAGAGTTATTAATGTCTCCAGTGTGGAGGTCGCCCTTGCAATCGTTGGAG
 GGGGCTATACTCCATCCAAATATGCAGTGAAGGTTTCAATGACAGCTTAAGACGGGACATGA
 AAGCTTTTGGTGTGCACGTCTCATGCATTGAACCAGGATTGTTCAAAACAACTTGGCAGATC
 CAGTAAAGGTAATTGAAAAAACTCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAAC
 AATATGGAGAAGGTTACATTGAAAAAGTCTAGACAACTGAAAGGCAATAAATCCTATGTGA
 ACATGGACCTCTCTCCGGTGGTAGAGTGCATGGACCACGCTCTAACAAGTCTCTTCCCTAAGA
 CTCATTATGCCGCTGGAAAAGATGCCAAAATTTTCTGGATACCTCTGTCTCACATGCCAGCAG
 CTTTGCAAGACTTTTTATTGTTGAAACAGAAAGCAGAGCTGGCTAATCCCAAGGCAGT**GTGAC**
 TCAGCTAACCACAAATGTCTCCTCCAGGCTATGAAATTGGCCGATTTCAAGAACACATCTCCT
 TTTCAACCCCATTCCTTATCTGCTCCAACCTGGACTCATTTAGATCGTGCTTATTTGGATTGC
 AAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGTCCCTGCTCAAGTTTTCTTTGAAAAG
 GAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCTGTATTTAGGCTTTGCCTGCTTGG
 TGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGATCTTTACCGTGGCCTGCCCCA
 TGCTTATGGTCCCCAGCATTTACAGTAACTTGGAATGTTAAGTATCATCTCTTATCTAAATA
 TTAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKG FHVIAACLTE
SGSTALKAETSERLRTVLLDVTDPENVKRTAQWVK NQVGEKGLWGLINNAGVPGVLAPTDWLT
LEDYREPIEVNLFGLISVTLNMLPLVKKAQGRVIN VSSVGGRLAIVGGGYTPSKYAVEGFNDS
LRRDMKAFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGE GYIEKSLDKLKG
NKS YVNMDLSPVVECMDHALTS LFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELAN
PKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGG
CGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGT
AAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATAC
AATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGA
AGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTA
CAAATTCCGTCGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTT
GCAGGAGCATTTTTCAAACCAAGACCTTGTTTTCTGCTATTAACACCAAGTATAATAACAGA
AAGCTGCTCTACTCATCGACTGGAACATTCTTATATAAACCTCAAAAAGGACTTTTTTCACAG
GGTACCTTTAGTGGTTGCCAATCTGGGCATGTCTGAACAACTGGGTTATAAACTGTATCAGG
TTCCTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTTGAAGA
AGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATT
AAAGAGTATATGCAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGT
AAACAGATTAAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGGCAGCAAGAGAGAA
GAACATCCAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCC
AAATTCTGAATTTCTTCATTTCATGTGTTATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAG
CTGTAACTACAACCACCATCTCGATGTAGTAGACAATCTGACCTTAATGGTAGAACACACTGA
CATTCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCATAAAGCCTTAGACTTAGA
TGACAGATGGCAATTCAAGAGATCTCGGTTGTTAGATACACAAGACAAACGATCTAAAGCAAA
TACTGGTAGTAGTAACCAAGATAAAGCATCCAAATGAGCAGCCCAGAAACAGATGAAGAAAT
TGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACATTT**TGAT**CCTTTTAACCTTA
CAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAACATTTCTATTGTTTTTACTATGTT
GAGCTACTTGCAAGTAAGTTCATTTGTTTTTACTATGTTACCTGTTTGCAGTAATACACAGAT
AACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCAGATGCTTTTATTTCCAAAC
CTTTTTTTCACCTTTCATAAGTTGTTGAGGGGAAGGCTTACACAGACACATTCTTTAGAATT
GGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACTTAGGGAAGACAAGTC
AGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATTGAGACCATGTCTA
TTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTATTTTCAAATATGGAAAGAAATTTATAT
GAAAATTTATCTGAGTCATTAAATTTCTCCTTAAGTGATACTTTTTTAGAAGTACATTATGGC
TAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCTAAATTT
AAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLEGEVKGEAKNSITDSQMDDVEVVYTIQK
YIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFS
NQDLVFLLLTPSIITESCSTHRLEHSYKPKGLFHRVPLVVANLGMSEQLGYKTVSGSCMST
GFSRAVQTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKR
EIEKRRGAQIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNH
HLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDLRWQFKRSRLDQTQDKRŠKANTGSSN
QDKASKMSSPETDEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCC
CAAGCAGCGCGCAGCGAACGCCCCGCCGCCGCCACACCCTCTGCGGTCCCCGCGGCGCCTGCCACCCTTCCCTCC
TTCCCCGCGTCCCCGCCTCGCCGGCCAGTCAGCTTGCCGGGTTCGCTGCCCCGCGAAACCCCGAGGTCACCAGCC
CGCGCCTCTGCTTCCCTGGGCCGCGCGCCGCCCTCCAGGCCCTCCTTCTCCCCTGGCCCGGCGCCTGGCACCAGGGG
ACCGTTGCCTGACGCGAGGCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCCTGCTCGCCTCTTCCACCAACT
CCAACTCCTTCTCCCTCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCCGCTGCCGTAGCGCCGC
TTCCCGTCCGGTCCCAAAGGTGGGAACGCGTCCGCCCGGCCCCGACCA**ATG**GCACGGTTTCGGCTTGCCCGCGCTT
CTCTGCACCCTGGCAGTGCTCAGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGA
CGTCTTTACGTGTCAAAGGCTTCAAACAAGACGATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATC
TGTCCCCAGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAA
AGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTCACGTTACAAGAAGTTTGATGAATTCTTC
AAAGAATACTTGAAAATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAA
AATTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAA
ATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCCTGGTGAACCTCCAGTACCACTTTACAGAT
GAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCTCGCAAATTGAAGCTC
CAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTTCGCTCAAGGCTTAGCGGTTGCGGGAGATGTGCTGAGCAAG
GTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAGATGATCTACTGCTCCCACTGCCGGGGT
CTCGTGAAGTGTGAAGCCATGTTACAATACTGCTCAAACATCATGAGAGGCTGTTTGGCCAACCAAGGGGATCTC
GATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGAGGGTCCTTTCAACATTGAA
TCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGATAATAGTGTTCAAGTGCT
CAGAAGGTTTTCCAGGGATGTGGACCCCCCAAGCCCTCCAGCTGGACGAATTTCTCGTTCATCTCTGAAAGT
GCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACCACAGCAGCTGGCACTAGTTTGGACCGA
CTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTCTGGTCCTCCCTTCCGAGCAACGTTTGCAAC
GATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTACCTGTTT
GCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACCAGACATA
CTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCA
GAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGTCTT
GGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCTGGTTATGCAGAGAGAGTGAGAT**TAAT**TCTCA
AACTCTGAGAAAAAGTGTTTCATCAAAAAGTTAAAGGCACCAAGTTATCACTTTTCTACCATCCTAGTGACTTTGC
TTTTTAAATGAATGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAAGTGCTGACTTTGTTTT
TCATTAGTTTTGGGAGGAAAAGGACTGTGCATTGAGTTGGTTTCTGCTCCCCCAAACCATGTTAAACGTGGCT
AACAGTGTAGGTACAGAACTATAGTTAGTTGTGATTTGTGATTTTATCACTCTATTATTTGTTTGTATGTTTT
TTCTCATTTTCGTTTGTGGGTTTTTTTTTCCAAGTGTGATCTCGCCTTGTTTCTTACAAGCAAACAGGGTCCCTT
CTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTTATCATGTTATC
TTATTAAAAGAAAAAGCCCAAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQG
 STCCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVK
 TYGHLYMQNSELFKDLFVELKRYVVGNNLEEMLNDFWARLLERMFRLVNSQYHFTDEYLEC
 VSKYTEQLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMI
 YCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPI
 DVKISDAIMNMQDNSVQVSQKVFQGC GPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAA
 GTSLDRLVTDVKEKLKQAKKFWSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGL
 ANQGNNEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFFDISDESSGEGSGSGCEY
 QQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:**Signal peptide:**

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTG
ACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGC
AACTTACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGC
CACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGG
ACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATT
GGTTCCTGAGAGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCC
CCTGTGATCATTTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAA
ACAAGCATTCCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGC
CTTTGTAGGAGCTCTGAGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAG
CACACCTACCAGACACTCTTCTTCTCCACCTCACTCTCCCACTGTACCCACCCCTAAATCAT
TCCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTTTGTTGCTCTCTCTAGTGTCTT
CTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAA
GATTCCAGGAACTGTAGCTTCCTAGCTAGTGTCAATTTAACCTTAAATGCAATCAGGAAAGTA
GCAAACAGAAGTCAATAAATATTTTTTAAATGTCAAAAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRK
FMTVSGLPKKQCPDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTC
AGAGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGTCTGCTTCTTACCC
TGCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCT
ACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCT
TCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCG
GAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATC
CCCCTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGT
TTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCT
GCACTCTGGTGTCTGTGCTCTGTGCAGAGCCCAAGGAAGGTCCTGCAGGAGGTCCGGAGAGTAC
TGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCT
TCATGTGGCAGCAAGTTTTCTGAGCCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCA
GAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCC
CTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCC
CAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGC
CTATCTATCTTCCACTGAGAGGGGACCTTAGCAGAATGAGAGAAGACATTCATGTACCACCTACT
AGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATCCCGCCTTCGACAGTGA
AAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCCTGTTGTATCCTCAACTG
CAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTCCCTTTCCTTCGTTCCCAT
GGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGTCTCTAGGAACTGGTCACAA
AAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCTCTCCCCACTACCACCTTCTT
CCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGATGCCAGAGCAAGACTCAAAGAG
GCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAAACCAG

FIGURE 28

MDILVPLLQLLVLLLTLPPLHLMALLGCWQPLCKSYFPYLMVLTTPKSNRKMESKKRELFSSQIK
GLTGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAP
GEDMRQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQ
VFEPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFSSKA
LICSFPSLQLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCCAAGCCCCTTTACCT**ATG**CTGCTGCTAACGCTGCTGCTG
CTGCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCCAGAAAGTCTCT
TCTGCCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTG
CCTCATCGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGG
GATGGC**TAA**GAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGC
CTTATTTAAAGTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTT
CATTAAGAAATTGTTCCCTTCCCCTGTGTTCAATGTTTGTAAGATTGTTCTGTGTAAATATG
TCTTTATAATAAACAGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAG
TTCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACT
CCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAATCA
TGTCGGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGA
TGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTATTTTGG
GATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCA
TAGAATTGGACACAGAAAGGGAAAATATGAAGTGCGTGCTGGGGTTTGTATCGTATCCACAG
GCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTG
AGCTTTTCCAAATCACAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCCAGCCACTGT
GGACATTTGCCATCCTCATTTTCTTCTGGGTCCCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAA
CTGCAGGAGCTGCCCAGGTTATGGAAGGCGGCCAAGTGGAATATAAGCCCCCTTTCGGGCATTC
GGTACATGTGGTTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCC
AGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTATTTCAACAGAAGTAAAAATGATCCTC
CTGATCATCCCATCCTTTTCGTCTCTCTCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGA
AAGGGTCATTTTTAATCTCTGTGGTGAGGATTCCGAGAATCATTGTCATGTACATGCAAAACG
CACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGATGCTGCTACTGCTGTT
TCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATATACTACAACCTGCTATTA
ATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAATCTTGTCCAAGAACTCAAGTC
ACTTTACATCTATTAACCTGCTTTGGAGACTTCATAATTTTTCTAGGAAAGGTGTTAGTGGTGT
GTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTACAATCGGGCATTCCAGGTGTGGGCAG
TCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGCCCATAGTTTTTTATCTGTGTTTG
AAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTTGATCTGGAAACAAATGATGGATCGT
CAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTAAAAGGAGCAACAAATTAA
ACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGAGGGAACAGAACTCCAGG
CCATTGTGAGAT**TAG**ATACCCATTTAGGTATCTGTACCTGGAAAACATTTCTTCTAAGAGCCA
TTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTTTTTAAAGACCTAA
TAAACCCTATTCTTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYDYTNDLS
IELDTERENMKCVLGFAIVSTGITAVLLVLI FVLRKRIKLTVELFQITNKAISSAPFLLFQPL
WTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILAC
QQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSLISVVRIPIRIVMYMQN
ALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKILSKNSS
HFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVF
ETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQ
AIVR

Important features:**Signal peptide:**

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

[illegible]

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCP
 AGCQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSL
 PRWRESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLA
 VTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTTSSQNRPRADPGIQRQDP
 SGAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQL
 LADVAQALDIGPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISF
 VTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVV
 EPNFANKAVCRTNGFYSLHVQSWFGLHKTLLQPLVKRVCDTDRLACSKTCLNSADIGFVIDGSS
 SVGTGNFRTVLQFVTNLTKEFEISDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGY
 WSGGTSTGAAINFALQLFKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAW
 AAQEELEVIATHPARDHSFFVDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:**Signal peptide:**

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
 431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAA
GAAATTGCCAAACCATGTCTTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTT
TAATTAAGCATGGAATACAGAAAACAACAAAAAACTTAAGCTTTAATTTTCATCTGGAATTCCA
CAGTTTTCTTAGCTCCCTGGACCCGGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCAC
GTGGTGCTCTCCGACTACTCACCCCGAGTGTAAGAACCTTCGGCTCGCGTGCTTCTGAGCTG
CTGTGGATGGCCTCGGCTCTCTGGACTGTCTTCCGAGTAGGATGTCACTGAGATCCCTCAA
TGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCTTCCCCAC
TACAATGTGATAGAACGCGTGAACCTGGATGTACTTCTATGAGTATGAGCCGATTTACAGACAA
GACTTTCACCTTCACACTTCGAGAGCATTCAAACCTGCTCTCATCAAAATCCATTTCTGGTCATT
CTGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAA
AAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAG
GAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGA
CAAGATTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTATGGCATTCAAGTGGGTGTA
ACTGAGTTTTGCCCCAATGCCAAGTACGTAATGAAGACAGACACTGATGTTTTTCATCAATACT
GGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTTTTTCACAGGTTATCCT
CTAATTGATAATTATTCCTATAGAGGATTTTACCACAAAACCCATATTTCTTACCAGGAGTAT
CCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATAATGTCCAGAGATTTGGTG
CCAAGGATCTATGAAATGATGGGTACGTAACCCATCAAGTTTGAAGATGTTTATGTCTGGG
ATCTGTTTGAATTTATTAAGGTGAACATTCATATTCAGAAAGACACAAATCTTTTCTTTCTA
TATAGAATCCATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCC
AAGGAGATCATCACTTTTTGGCAGGTCATGCTAAGGAACACCACATGCCATTATTAACTTCAC
ATTCTACAAAAGCCTAGAAGGACAGGATACCTTGTGGAAAGTGTTAAATAAAGTAGGTACTG
TGGAAATTCATGGGGAGGTCAGTGTGCTGGCTTACACTGAACTGAACTCATGAAAAACCCA
GACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGGCCCTTCAAAGATGATATGTGG
AGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGGACCAAACAATTTG
GACATGTCAATTCTGTAGACTAGAATTTCTTAAAGGGTGTTACTGAGTTATAAGCTCACTAGG
CTGTAAAAACAAAACAATGTAGAGTTTTATTTATTGAACAATGTAGTCACTTGAAGGTTTTGT
GTATATCTTATGTGGATTACCAATTTAAAAATATATGTAGTCTGTGTCAAAAACTTCTTCA
CTGAAGTTATACTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTT
GCAGTATTTACAGTTATTATTATTTAAAATTACTTCAACTTTGTGTTTTTAAATGTTTTGAC
GATTTCAATACAAGATAAAAAGGATAGTGAATCATTCTTTACATGCAAACATTTTCCAGTTAC
TTAACTGATCAGTTTATTATTGATACATCACTCCATTAATGTAAAGTCATAGGTCATTATTGC
ATATCAGTAATCTCTTGGACTTTGTAAATATTTTACTGTGGTAATATAGAGAAGAATTAAAG
CAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDF
HFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EKED
KMLALSLEDEHLLYGDIIRQDFLDTYNNLTTLKTIMAFRWVTEFCPNAKYVMKTDTDVFIN TGN
LVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPR
IYEMMGHVKPIKFEDVYVGICLNLLKVNHIHPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKE
IITFWQVMLRNTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVG
 YTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLTDDFYVKGIFYCAECRAGWYGGDCM
 RCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDN
 RDGQIIKRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVL
 DKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFC
 NNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMPVQVSRETPLHQLYSAA
 FSKQKLQSAPTKKPPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCCLRTGKWSGRAPS
 CIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAAH
 CVTDLGKVTMIKTADLKVVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLL
 DKARISTRVQPICLAASRDLSTSFQESHITVAGWNVLADVRS PGFKNDTLRSGVVSVDLLC
 EEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMGLVSWSY
 DKTCSHRLSTAFTKVLPFKDWIERNMK

Important features of the protein:**Signal peptide:**

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
 474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACGTGGCTTA
ATCTGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTT
GGCTGGTTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGC
TTCTGTTGCTGGTCTTGCCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTGTATT
CAGAACTCTGTAAAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAG
ACGGCTGTGCGAGCCTCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAG
ACGAGCCTGGCCTAGACAACCCTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCAGTGGACT
CTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAA
ATCGAGCTTTGAGTGTCTTTCGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAA
ATTCTGAAAACACCCTGCCCCCTGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAAATTACCAGCA
TCAAGATCAATCGAGTAGATCCCAGTGAAAGCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCC
ATATCATTATCCAACACATTTATCGTGATGGGGTGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTC
TAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACCTACGCTGTGCGTCTCCTGCGGCAGCCCTGCCAGG
TGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCAGCAGGAACAATGGACAGGCCCCGGATGCCTACAGAC
CCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGGGAATAAACTGGTGC
AGGTGGATGAGCCTGGGGTTCATCTTCAATGTGCTGGATGGCGGTGTGGCATATCGACATGGTCAGCTTGAGG
AGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCAGAAAGTGGCGCTCATCTGATTC
AGGCCAGTGAAAGACGTGTTACCTCGTCGTGTCCCGCCAGGTTCCGCAGCGGAGCCCTGACATCTTTCAGGAAG
CCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACTCCCAAGCCCCCTCCATCCTA
CAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCGGCATGACCGTCGCAGGGG
GAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTGAGCCCGGAGGAGTCATAAGCAGAGATG
GAAGAATAAAACAGGTGACATTTTGTGTAATGTGGATGGGGTCGAACCTGACAGAGGTGAGCCGGAGTGAGGCAG
TGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGCCCCAGGAAG
ACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCCTGGGTCA
TGTGGCTGGAATTACCACGGTGCTTGATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTCTGG
GCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAAACAAACCTTTTTTTCATCAAATCCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTGCTGTCAATGGTAGAAGTACATCAGGAA
TGATACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTG
GCACTTTTTTATAGAATCAATGATGGGTGAGAGGAAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAAACACT
ATATTTATCTTGTGAGTTTTTATATTTAAAGAAAGAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAA
TGAAAGCCAGTTACACCTCAGAAAAATATGATTCCAAAAAATTAACCTACTAGTTTTTTTTTCACTGTGGAGGAT
TTCTCATTACTCTACAACATTGTTTATATTTTTTCTATTCAATAAAAAGCCCTAAAACAATAAATGATTGATT
TGTATACCCCACTGAATTCAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTAT
GGCCATTTTTTAATTTACAGCTAAAATATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTTCATCAAACAAGAAT
AAATATTTTTTCAGAAGTTAAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATA
 PSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFK
 KINRALSVLRRTKSGSAVANHADQGRESENTTAPEVFPRLYHLIPDGEITSIKINRVDPSES
 LSIRLVGGSETPLVHIIIIQHIIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPC
 QVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVKRVDEPGVFIFNV
 LDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQVRQSPDIFQE
 AGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPYIV
 ISVEPGGVISRDGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQED
 CSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGN
 KPFFIKSIVEGTPAYNDGRIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:**Signal peptide:**

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
 467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAA
GCTTTCTTGCCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGGCTT
CAACCTGACTTTCCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCTGAC
AGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCTTAAAGCAAA
GGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGA
AGCATCCACGAAGAAGGTAGAACTTGACAACCTGTCCTTCTGTGTCTCCTTACCTCAGAGGCCA
GAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATCCCAAAGT
GTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATCCTCGTTCC
CCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTTCCTGCAGAG
GCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGTTTAATCGAGC
CAAACCTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGACTGCTTTATATT
CCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAGCATCCCAA
GCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGGGGG
TGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACCTACTGGGG
ATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCG
GCCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCACACTAGAGACAAAGGCAATGAGGT
GAACGCAGAACGGATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGGTTGAG
TAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGA
TTTCTGGTTTGGTGCATGACCCTGGATCTTTTGGTGATGTTTGAAGAACTGATTCTTTGTTT
GCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTAAGAACCTGTTACAGCTCATT
GTTGAGCTGAATTTTTCCTTTTGTATTTTCTTAGCAGAGCTCCTGGTGATGTAGAGTATAAA
ACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGTATCATGAGGGTTAAATATTGTAATATGG
ATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAAAATGAACGCTATTTGAGGACT
CTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGATAAAAGGCCACAGGAAA
TAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAAAGGTACGAAGA
TACAATACTGTTATTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGAGGTGAGAAG
GCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGAG
GTAGCAGGAGGGTGGAGTGTGCGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGA
TAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCTTCAGTGATGCCACCAGAGAATACATTC
TCTATTAGTTTTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTAGCAGT
TTACAAGTTTACATATTAATAATAATAATATGTCTATCAAATACCTCTGTAGTAAAATGTG
AAAAAGCAAAA

FIGURE 42

MGFNLT FHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTL
TNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAI
LVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKGKFNRAKLLNVGYLEALKEENWDC
FIFHDVDLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNN
YWGWWGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTD
GLSSCSYKLVSVEHNPLYINITVDFWFGA

Important features:**Signal peptide:**

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACG**ATG**GCACTGAGCTCCCAGATCTG
GGCCGCTTGCCTCCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCACA
ACAGACGGGACAACCTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGAT
GCCCATGTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGG
CTGCTGTCATCGATCAAAGTGTGGGATGTGCTGCAAGACG**TAG**AACCTACCTGCCCTGCCCCC
GTCCCCCTCCCTTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAATGGCTG
GTTCTTTTGTTTTCCAAA
AAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183

FIGURE 44

MALSSQIWAACLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHF
PICIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCAC
CCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGT
CGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTAT
TGTCTGGACCTTCAACACAACCCCTCTTGTCAACATACAGCCAGAAGGGGGCACTATCATAGT
GACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAG
CAAAGTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCA
GCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCAT
GGGCTGTCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGG
GGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGG
GTCCATCCTCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAG
GAACCCTGTCAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGC
TGATGACCCAGATTCCCTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCCCTCCTGCTCAGTCT
CTTTGTACTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGA
GAAGAAGAGAGTGGACATTTGTGCGGAAACTCCTAACATATGCCCCCATCTGGAGAGAACAC
AGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGATCCAGCAAATACGGT
TACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCACTGCTCACGATGCCAGA
CACACCAAGGCTATTTGCCTATGAGAATGTTATCTAGACAGCAGTGCACTCCCCTAAGTCTCT
GCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTI
QPEGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYE
HLSKPKVTMGLQSNKNGTCVTNLTCCMEHGEEVDVIYTWKALGQAANESHNGSILPISWRWGES
DMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKR
ERQEYIEEKRVDICRETPNICPHSGENTHEYDTIPHTNRTILKEDPANTVYSTVEIPKKMEN
PHSLLTMPDTPRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAAT
GGATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTC
AGCTTAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCCA
GGAATTATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAA
AGAGCGTGCTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTC
ATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGT
AATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCAT
CCAGAATCCTTCAACTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTTCAAT
AAACCCACCAGTAACGACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCT
GAAGAAAACAAACATAGGCTTATCCACTTCTCAGTATTTTATAGGTCTATTGCTTGTGGAATT
CTGGAGGTCCCTGTTTGGGCTCAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTC
TCTAAGCGAAGAAGTCAAATTGTGTAGTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGA
AAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAI
PATMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANC
EFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFHDSEENKHRLIHF
SVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGT
GAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGA
GCAGATCCGTGGGCTGCAGACCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCTCGAA
CTGTGACATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGC
CAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAG
CGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATG
CAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACATCATCAC
TCCAGGCTCTGCCACTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAAC
ACTGGCCCCCAGCACCTCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGG
GCAGGCTGTTAGGCCCTTTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACC
CCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAAGIAAVLSGKCKY
KSSQKQHSPVPEKAIPILITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

[illegible]

FIGURE 52

MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAA
 GSKVSEALGQGTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGAD
 AVRGSWQGVPGHSGAWETSGGHGIFGSQGGGLGGQGGQGNPGGLGTPWVHGYPGNSAGSFGMNPQ
 GAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGSGSQS
 GSSGSGSNGDNNNGSSSGSSSGSSSGSSSGSSSGSSSGSSSGNSGGSRGDSGSESSWGSSTG
 SSSGNHGGSGGGNGHKPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRLLGGSGDN
 YRQGSSSWGSGGGDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDAINKDQRSSRI

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,
 74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,
 155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205,
 218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252,
 249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281,
 279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298,
 295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,
 323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387,
 383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCT
GGCTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCT
GGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAAGTCCCGCCGGCTCCAGT
GTTTCCCACAGCCCCCAAACGGAAGTGGTTTTGGGGTACCTGGGCCTGATCACTCCTACAG
AGGAGGGGCTTGAAGGACTCGACCCAGATGTGCGCCACCTATTCCCAGGGCTTTACGGTATGGC
TGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATG
CCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGGAG
AAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCCGCT
TCCATTTCAACATCCTGAAGTCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTG
ACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCC
TCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGC
CCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAAGAGAAAAGAAGCCAGCATA
TCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCT
GCCGCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGCGTGCACCCCTCCCCACTC
AGGGTATTGATGATTTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGGATTTTATTGATGTGC
TTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTGATGAGGATATAAGAGCAGAGGCTG
ACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTCCTGGGTCTGTACAACC
TTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGCAAGAGCTTCTGAAGGACC
GCGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCCTTCTGACCATGTGCGTGA
AGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCGATGCTGCACCCAGGACATTG
TTCTCCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCCTCATCGATATTATAGGGGTCC
ATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACCCCTTCCGCTTTGACCCAGAGA
ACAGCAAGGGGAGGTACCTCTGGCTTTTATTCTTTCTCCGCAGGGCCCAGGAACTGCATCG
GGCAGGCGTTGCCATGGCGGAGATGAAAGTGGTCTGGCGTTGATGCTGCTGCACTTCCGGT
TCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGCGCGCCGAGGGCGGGC
TTTGGCTGCGGGTGGAGCCCCCTGAATGTAGGCTTGCACTTTCTGACCCATCCACCTGTT
TTTTTGAGATTGTCATGAATAAAACGGTGCTGTCAAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWGH
LGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIR
FLKPWLGE GILLSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRL
DMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHD
GRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFKDKAKSKTLDFIDVLLLSKDEDGKALSD
EDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQL
PFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYD
PFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLEL
IMRAEGGLWLRVEPLNVGLQ

Important features:**Transmembrane domains:**

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPF
ARDAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGGCAGTGGAGCACCCAGCAGGCCGCCAAC**ATG**CTCTGTCTGTGC
 CTGTACGTGCCGGTCATCGGGGAAGCCCAGACCGAGTTCAGTACTTTGAGTCGAAGGGGGCTCCCTGCCGAGCTG
 AAGTCCATTTTCAAGCTCAGTGTCTTCATCCCCCTCCAGGAATTCCTCCACCTACCGCCAGTGGGAAGCAGAAAATT
 GTACAAGCTGGAGATAAGGACCTTGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAG
 AAGAAGCTGAGGCTGGTGTAAAGATTTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATGCGAG
 TCCCTGCGGGACTTGGGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCTCAAGAGCATGGATAAAAAACGGC
 ACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAAACATCCCCGAGATCATC
 CTCTACTGGAAGCATTCCACGATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAG
 AGGCAGACGGGGATGTGGTGGAGACACCTGGTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCC
 CCCCTGGACAGGCTCAAGGTGCTCATGCAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTC
 ACTCAGATGATTCGAGAAGGAGGGGCCAGGTCACTCTGGCGGGGCAATGGCATCAACGTCCTCAAAATTGCCCCC
 GAATCAGCCATCAAATTCATGGCCTATGAGCAGATCAAGCGCCTTGTGGTAGTGACCAGGAGACTCTGAGGATT
 CACGAGAGGCTTGTGGCAGGGTCCTTGGCAGGGGCCATCGCCCCAGAGCAGCATCTACCAATGGAGGTCCTGAAG
 ACCCGGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGAATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAG
 GGGGTGGCCGCCTTCTACAAAGGCTATGTCCCAACATGCTGGGCATCATCCCCTATGCCGGCATCGACCTTGCA
 GTCTACGAGACGCTCAAGAATGCCTGGCTGCAGCACTATGCAGTGAACAGCGCGGACCCGGCGTGTGTGTGCTC
 CTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGCTACCCCTGGCCCTAGTCAGGACCCGGATG
 CAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTCTTCAAACATATCCTGCGGACCGAG
 GGGCCTTCGGGCTGTACAGGGGGCTGGCCCCAACTTCATGAAGGTCACTCCAGCTGTGAGCATCAGTACGTG
 GTCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTGCAGG**TGA**CGGGGGGAGGGCCCGCCGGCAGTGGACTCG
 CTGATCCTGGGCCGACGCTGGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACACTAAGCTGTCTCGAGCC
 AAGCTGTGAAAACCTTAGACGCACCCGACGGGAGGGTGGGGAGAGCTGGCAGGCCAGGGCTTGTCTGTGACC
 CCAGCAGACCTCCTGTGTGGTTCAGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTCAGCAGGCTCCGGGCT
 CACATGTGTAAGGACAGGACATTTCTGTCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCTTAGTTC
 TTCCATTTACCCCTTGACGCCAGCTGTTGGCCACGGCCCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGCCCT
 CTTGCTGCCTGCCTGTCTGCTGAGGTAAGGTGGGAGGAGGGCTACAGCCACATCCCACCCCTCGTCCAATCCC
 ATAATCCATGATGAAAGGTGAGGTACGTGGCCTCCCAGGCCTGACTTCCCAACCTACAGCATTGACGCCAACTT
 GGCTGTGAAGGAAGAGGAAAGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGTGATGGCTGGGGCTCTCG
 GGCATGCTTGGGAGTGCAGGGGGCTCGGGTGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTG
 CTCTGAGCTGGCCTGGACCTGTGAGGATGGGGCCCCACCTCAGAACCAAACTCACTGTCCCCTGTGGCATGAG
 GGCAGTGGAGCACCATGTTTGGGGCGAAGGGCAGAGCGTTTGTGTGTTCTGGGGAGGGAAGGAAAGGTGTTGG
 AGGCCTTAATTATGACTGTTGGGAAAAGGGTTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGC
 TTCCAGAGGAAGACGAGGGAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCTGTGC
 CAACCCAGCAGGGGCGCAGCGGGACCAGCCCCACATTCCACTTGTGTCACTGCTTGGAACTATTTATTTTGTA
 TTTTATTTGAACAGAGTTATGTCCTAACTATTTTATAGATTTGTTTAATTAATAGCTTGTCAATTTCAAGTTTAT
 TTTTATTTTATATTTATGTTTATGTTGATTGTACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGGAGAA
 GGGGGGCTTGGGGCGCTGCAGTCACTGTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAG
 AAGGCAGCAGCCCTGGCTCCTTTTCTTTGGCAGGTTGGGGAAGGGCTTGGCCCCAGCCTTAGGATTTAGGGTTT
 GACTGGGGGCGTGGAGAGAGAGGGAGGAACCTCAATAACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCGA
 GGGTTTCTTTATTTTCACTCTTTTCTGAATGTCAAGGCAGTGAGGTGCCTCTCACTGTGAATTTGTGGTGGGCGGG
 GGCTGGAGGAGAGGGTGGGGGGCTGGCTCCGTCCCTCCAGCCTTCTGCTGCCCTTGCTTAACAATGCCGGCCAA
 CTGGCGACCTCACGGTTGCACTTCCATTCCACCAGAATGACCTGATGAGGAAATCTTCAATAGGATGCAAAGATC
 AATGCAAAAATTGTTATATATGAACATATAACTGGAGTCGTCAAAAAGCAAATTAAGAAAAGAAATTGGACGTTAGA
 AGTTGTCATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDL
DGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEEKILKSM
DKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTPDEFTVEERQTGMWWRH
LVAGGGAGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLK
IAPESAIKFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQSSIYPMEVLKTRMALRKTGQ
YSGMLDCARRILAREGVAAFYKGYVPNMLGIIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVF
VLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAP
NFMKVIPAVSISYVVYENLKITLGVQSR

Important features:**Signal peptide:**

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCC**ATG**
GCTTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCA
ATTGCACTCATCATTGGCTTTGGTATTTAGGGAGACACTCCATCACAGTCACTACTGTGCGC
TCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAAACCT
TCTGATATCGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAA
GGCAAAGATGAGCTGTCGGAGCAGGATGAAATGTTTCAAGAGGCCGGACAGCAGTGTTTGCTGAT
CAAGTGATAGTTGGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACC
TACAAATGTTATATCATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGA
GCCTTCAGCATGCCGGAAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAG
GCTCCCCGATGGTTCCCCCAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAAC
TTCTCGGAAGTCTCCAATACCAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTG
TCTGTGCTCTACAATGTTACGATCAACAACACATACTCCTGTATGATTGAAAATGACATTGCC
AAAGCAACAGGGGATATCAAAGTGACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTG
CTAAACTCAAAGGCTTCTCTGTGTGTCTCTTCTTTCTTTGCCATCAGCTGGGCACCTTCTGCCT
CTCAGCCCTTACCTGATGCTAAAA**TAA**TGTGCCTTGCCACAAAAAGCATGCAAAGTCATTG
TTACAACAGGGATCTACAGAACTATTTACCACCAGATATGACCTAGTTTTATATTTCTGGGA
GGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAGAGCAAGAAACAAAAAGAAGCC
AAAAGCAGAAGGCTCCAATATGAACAAGATAAATCTATCTTCAAAGACATATTAGAAGTTGGG
AAAATAATTCATGTGAACTAGACAAGTGTGTTAAGAGTGATAAGTAAAATGCACGTGGAGACA
AGTGATCCCCAGATCTCAGGGACCTCCCCCTGCCTGTCACCTGGGGAGTGAGAGGACAGGAT
AGTGATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGCTGTAATGTTGCTCTGAGGAAGCC
CCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCACAAATTAAGCTGTAGTATGTAC
CCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCGGCTGCATTTTAGTAATGGG
TCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTCTCTTCCCAACTGACAAA
TGCCAAAGTTGAGAAAAATGATCATAATTTAGCATAAACAGAGCAGTCGGGGACACCGATTT
TATAAATAAACTGAGCACCTTCTTTTTAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIK
LSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAG
TYKCYIITSKGKGNANLEYKTGAFSMPEVNVNDYNASSETLRCEAPRWFPQPTVVWASQVDQGA
NFSEVSNTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTSEIKRRSHLQ
LLNSKASLCVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCA**ATG**GCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACA
AGCACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGAGGGGCAGTATGG
TAGTGGGCTACCCCCTGGTGGTGGTTATGGGGGTCTGCCCCTGGAGGGCCTTATGGACCACC
AGCTGGTGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCTCTGGAACTCCAGGAGG
ACCATATGGCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGC
CCAGCAGCCTGGGCTTTATGGACAGGGTGGCGCCCCCTCCCAATGTGGATCCTGAGGCCTACTC
CTGGTTCCAGTCGGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGC
CCTGGTCAACTGCAATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTT
TGACAAGACCAAGTCAGGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCA
GCAGTGGAAGAACCTCTTCCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGA
GCTGCAGCAAGCTCTGTCCCAAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTTCTGGT
CTCCCGCTACTGCCCACGCTCTGCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTG
CACCCAGCTGCAGGTGCTGACAGAGGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACAT
CCGGCTCAGCTTCGAGGACTTCGTCACCATGACAGCTTCTCGGATGCTAT**TGA**CCCAACCATCT
GTGGAGAGTGGAGTGCACCAGGGACCTTTCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACA
TCTCTTCTTTTCTGTCCCTCTAGAAGAACATTCTCCCTTGCTTGATGCAACACTGTTCCAAA
AGAGGGTGGAGAGTCCTGCATCATAGCCACCAAATAGTGAGGACCGGGGCTGAGGCCACACAG
ATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGAATGTCCTGATGGCCATGAGCAGTTGAGTG
GCACAGCCTGGCACCAGGAGCAGGTCTTGTAATGGAGTTAGTGTCCAGTCAGCTGAGCTCCA
CCCTGATGCCAGTGGTGAAGTTCATCGGCCTGTTACCGTTAGTACCTGTGTTCCCTCACCAG
GCCATCCTGTCAAACGAGCCCATTCTCTCAAAGTGGAATCTGACCAAGCATGAGAGAGATCT
GTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAATCCTTGTGTGTTAACTTCTAGC
TGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCATCTTTGGCCAGGCTTCTGCC
CCCTGCAGCTGGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCTTCAGTCTCCAGGAGACA
GTGGTCACCTCTCCCTGCCAATACTTTTTTTAATTTGCATTTTTTTTCATTTGGGGCCAAAAG
TCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGP
YGHFNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDP EAYSWFQSV
DSDHSGYISMKELKQALVNCNWSSFND ETC LM MINMFDKTKSGRIDVYGFSALWKFIQQWKNL
FQQYDRDRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQV
LTEAFREKDTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCA
TCTCCCAGGCCTCTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGG
GACCAACTTGCCTCAGCTCGGACAACCTTCTCCTCACTGGCCCCCTCTAACTCTGAACATCCGCA
GCCCCTCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCCTCTGAAGCTCAGCGTGCCTCC
ATCAGATGGCTTCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCT
GCCTGCCATGGATTCTGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGA
CCGCCTGGGGGAAGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCC
GGGCAGTGGCCCTTTGCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTC
ACTCCTCCACCAGGACTCGGAGTCCAGACGACTGCCCCGTTCTAATTCAGTGGGAGCCGGGGG
AAAAATCCTTTCCCAACGCCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTG
GGGTACCCTGAATCCCAGTGTGTCTGCGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAG
GCCCATGCCACACCCTGAGGGAATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGG
AAATATTAATCGGTATCCAGGAGGCAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTG
GGGGAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTCATCTATACCCAGGTATCAA
TAACCCATTTCTCCTGGAGTTCTCCGCCCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTT
CCCTAATCCTCCAAGCCCTAGGTTGCAGTGGGGCTAGAGCACGATAGAGGGAAACCCAACATT
GGGAGTTAGAGTCCTGCTCCCGCCCCCTTGCTGTGTGGGCTCAATCCAGGCCCTGTAAACATGT
TTCCAGCACTATCCCCACTTTTCAGTGCCTCCCCTGCTCATCTCCAATAAAATAAAAGCACTT
ATGAAA
AA

FIGURE 64

MQGRVAGSCAPLG LLLVCLHLPGLFARSIGVVEEKVSQNF GTNLPQLGQPSSTGPSNSEHPQP
ALDPRSNDLARVPLKLSVPPSDGFPFAGGS AVQRWPPSWGLPAMDSWPPEDPWQMMAAAAEDR
LGEALPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLLPRSNSLGAGGK
ILSQRPPWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPGEGIWGINNQPPGT SWGN
INRYPGGSWGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFP
NPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACT**ATG**GGG
TCTGGGCTGCCCCCTTGTCCCTCCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGT
ATGACTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTC
CTGGAATTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTC
CACCATGCAAGATCTCAACACCATGTTGTCTGCAACACAT**TGA**CAGCCATTGAAGCCTGTGTCC
TTCTTGGCCCGGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAG
GCCCCCACCTCCTGAGTGGCAATAAATAAAATTCGGTATGCTG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
21

FIGURE 66

MGSGPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSGTSV
TLHHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGG
GCCAGGTGCCCCGTCGCAGGTGCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAG
CCCCTTCCTCGGCGCTGCCAACCCGCCACCCAGCCC**ATG**GCGAACCCCGGGCTGGGGCTGCTT
CTGGCGCTGGGCCTGCCGTTCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACC
ACTTCTGCAAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAAC
CTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTG
GCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTAC
CGGCCCAGTAGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCTCAGGACTCCAAG
GAGACGGTGCAGGGCTGCCTGCCCATC**TAG**GTCCTCTCCTGCATCTGTCTCCCTTCATTGC
TGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAG
CAGGGAAGAAGGTACTTCAAAGACTCTGCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTT
TTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSTSSSSSDGNLRPEAITAIIVV
FSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

FIGURE 70

MGLFRGFVFLVLCLLHQSNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTASTYL
 FEATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGE
 KGEYIHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCSA
 GISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVQTEKASIMFMQSIDSVEFCNE
 KTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSG
 SMGGKDRNLNRMNQAAKHFLQLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTY
 PLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRA
 ADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMN
 DTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAY
 NLQAKANPETLTITVTSRAANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVT
 AFIESQNGHTEVLELLDNGAGADSFKNQDVYSRYFTAYTENGRYSLKVRAHGGANTARLKLRP
 PLNRAAYIPGWVNGEIEANPPRPEIDEDTQTTLEDFSRASGGAFVVSQVPSLPLPDQYPPS
 QITDLDATVHEDKIIILTWTAPGDNFQVGVQRYIIRISASILDLRDSFDDALQVNTTDLSPKE
 ANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPT
 PTPTPDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAAACCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCGGGCAGGGG
 TGACAACAGGTGTCATCTTTTTGATCTCGTGTGTGGCTGCCCTTCTTCAAGGAAAGACGCCAAGGTAATTTT
 GACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCAAGTTATGCCAGGATTTACTAG
 AGAGTGTCAACTCAACCAGCAAGCGGCTCCTTCGGCTTAACTTGTGGTTGGAGGAGAGAACCCTTTGTGGGGCTGC
 GTTCTCTTAGCAGTGTCTCAGAAGTGAAGTGTGCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCCCTCTTGCTGTTG
 GCTGCACATCAGGAAGGCTGTGATGGGAATGAAGGTGAAAACCTTGAGATTTCACTTCAGTCATTGCTTCTGCCT
 GCAAGATCATCTTTTAAAGTAGAGAAGCTGCTCTGTGTGGTAACTCCAAGAGGCAGAAGCTCGTTCTAGAA
 GGAAATGGATGCAAGCAGCTCCGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTG
 GGGGCCCCGGCTTTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGGTTCGCCGGGGG
 CTGCTTGCGTGGATTTCCCGGGTGGTGGTTTTGCTGGTGCTCCTCTGCTGTGCTATCTCTGTCTGTACATGTTG
 GCCTGCACCCCAAAGGTGACGAGGAGCAGCTGGCACTGCCAGGGCCAACAGCCCCACGGGGAAGGAGGGGTAC
 CAGGCCGTCCCTCAGGAGTGGGAGGAGCAGCACCAGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTC
 AAGGAGGAGCTGCAGGAGAGGAGTGAAGCAGCTCAGGAATGGGAGTACCAAGCCAGCGATGCTGGCCTGGGT
 CTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTCCTGGCCTTCTGCACTCGCAGGTGGACAAAGGAGAG
 GTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGCAGCAGTGCCTTTCGATAGCTTTACTCTACAGAAGGTGTAC
 CAGCTGGAGACTGGCCTTACCCGCCACCCCGAGGAGAAGCCTGTGAGGAAGGACAAGCGGGATGAGTTGGTGGAA
 GCCATTGAATCAGCCTTGAGACCCCTGAACAATCCTGCAGAGAAGCAGCCCCAATCACCCTGCTTACACGGCCTCT
 GATTTTCATAGAAGGATCTACCGAACAGAAAGGGACAAAGGACATTGTATGAGCTCACCTTCAAAGGGGACCAC
 AAACACGAATTCAAACGGCTCATCTTATTTTCGACCATTCAGCCCCATCATGAAAGTGAAAAATGAAAAGCTCAAC
 ATGGCCAACACGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTCGGCAGTTTCATGCAGAAT
 TTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGGAAAGAAGAAATAAT
 GAAGTCAAAGGAATACTTGAAAACACTTCCAAAGCTGCCAATTTCAGGAACCTTACCTTCATCCAGCTGAATGGA
 GAATTTTCTCGGGGAAGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGTCTCTCTTTTCTGT
 GATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAGGTATTT
 TATCCAGTTCTTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCACTCCCTCCCTTGGAAACAG
 CAGCTGGTCATAAAGAAGGAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTGAGTATCGGTGAGAC
 TTCATCAATATAGGTGGGTTTGTCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTAT
 CTCCACAGCAACCTCATAGTGGTACGGACGCTGTGCGAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATG
 GACGAGCTGACCCCCGAGCAGTACAAGATGTGCATGCAAGGCTCAAGGCCATGAACGAGGCATCCCACGGCCAGCTG
 GGCATGCTGGTGTTCAGGCACGAGATAGAGGCTCACCTTCGCAAAACAGAAACAGAAGCAAGTAGCAAAAAACA
TGAACTCCCAGAGAAGGATTGTGGGAGACACTTTTTCTTTCTTTTGCATTACTGAAAGTGGCTGCAACAGAGA
 AAAGACTTCCATAAAGGACGACAAAAGAATTGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTCTGT
 TGGGCTTTTTTACAACAGAAATCAAATCTCCGCTTTGCGCTGCAAAAGTAACCCAGTTGCACCCTGTGAAGTGTCT
 GACAAAGGCAGAATGCTTGTGAGATTATAAGCCTAATGGTGTGGAGGTTTGTATGGTGTGTTACAATACACTGAGA
 CCTGTTGTTTTGTGTGCTCATTGAAATATTCTGATTTAAGAGCAGTTTTGTAAAAAATTCATTAGCATGAAAGG
 CAAGCATATTTCTCCTCATATGAATGAGCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAAATATCAGAAGG
 CAGGAGAGGAGATAGGCTTATTATGATACTAGTGAGTACATTAAGTAAAATAAAATGGACCAGAAAAGAAAAGAA
 ACCATAAATATCGTGTATATTTTCCCAAGATTAACCAAAAAATAATCTGCTTATCTTTTTGGTTGTCTTTTAA
 CTGCTCTCCGTTTTTTTTCTTTTATTTAAAAATGCACTTTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTAC
 CACTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTACATTTTTATATTTTTTAAGAAGATACTTTGAGATGCA
 TTATGAGAACCTTCAGTTCAAAGCATCAAATTGATGCCATATCCAAGGACATGCCAATGCTGATTCTGTGAGGC
 ACTGAATGTCAGGCATTGAGACATAGGGAAGGAATGGTTTGTACTAATACAGACGTACAGATACTTTCTCTGAAG
 AGTATTTTCGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGACACTTTCTGCTTTACAGAAAAGGAAACT
 CATTACAGCTGGTGATATCGTGATGTACCTAAAAGTCAGAAACCACATTTTCTCCTCAGAAGTAGGGACCGCTTT
 CTTACCTGTTTAAATAAACCAAAGTATACCGTGTGAACCAACAATCTCTTTTCAAACAGGGTGCTCCTCCTGG
 CTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATATATATATATATTTGTGAAAGATCAATCCATCTG
 CCAGAATCTAGTGGGATGGAAGTTTTTGTCTACATGTTATCCACCCAGGCCAGGTGGAAGTAACTGAATTTTTT
 TAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAAATGCAATTTTATACCATTTCAAACATTTTTT
 TAAAAATAAATACAGTTAACATAGAGTGGTTTTCTTCATTGATGTGAAAATTATTAGCCAGCACCAGATGCTGAG
 CTAATTATCTCTTTGAGTCTTGTCTGTTTGTCTCACAGTAAACTCATTGTTTAAAGCTTCAAGAACATTCAA
 GCTGTTGGTGTGTTAAAAAATGCATTGTATTGATTTGTACTGGTAGTTTATGAAATTTAATTAAACACAGGCCA
 TGAATGGAAGGTGGTATTGCACAGCTAATAAAATATGATTTGTGGATATGAA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQALPRANSPTGKEGYQAVLQE
WEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHS
QVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPPEKPVRKDKRDELVEAIESALE
TLNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVK
NEKLNMAANTLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILEN
TSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPG
KKVFYPVLFSQYNPGIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDL
DIKGWGGEDVHLYRKYLHSNLIVVRTPVRLFHLWHEKRCMDELTPEQYKMCMQSKAMNEASH
GQLGMLVFRHEIEAHLRKQKQKTSSKKT

Important features:**Signal peptide:**

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCCTGGGGATC
CAGAAACCCATGATACCCTACTGAACACCGAATCCCCCTGGAAGCCCACAGAGACAGAGACAGC
AAGAGAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTCTCA
CTCCTCCCTCCCTCTCTCTCTGCTGTCTAGTCCTCTAGTCCTCAAATTCCCAGTCCCCTGC
ACCCCTTCCTGGGACACTATGTTGTTCTCCGCCCTCCTGCTGGAGGTGATTTGGATCCTGGCT
GCAGATGGGGGTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTGGCCAGCCTCT
TACCCTGAGTGTGGAAACAATGCCCAGTCGCCCATCGATATTCAGACAGACAGTGTGACATTT
GACCCTGATTTGCCTGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCCAGCCTTTGGAC
CTGCACAACAATGGCCACACAGTGCAACTCTCTCTGCCCTCTACCCTGTATCTGGGTGGACTT
CCCCGAAAATATGTAGCTGCCCAGCTCCACCTGCACTGGGGTCAGAAAGGATCCCCAGGGGGG
TCAGAACACCAGATCAACAGTGAAGCCACATTTGCAGAGCTCCACATTGTACATTATGACTCT
GATTCCTATGACAGCTTGAGTGAGGCTGCTGAGAGGCCTCAGGGCCTGGCTGTCCTGGGCATC
CTAATTGAGGTGGGTGAGACTAAGAATATAGCTTATGAACACATTCTGAGTCACTTGCATGAA
GTCAGGCATAAAGATCAGAAGACCTCAGTGCCTCCCTTCAACCTAAGAGAGCTGCTCCCCAAA
CAGCTGGGGCAGTACTTCCGCTACAATGGCTCGCTCACAACCTCCCCCTTGCTACCAGAGTGTG
CTCTGGACAGTTTTTTTATAGAAGGTCCAGATTTCAATGGAACAGCTGGAAAAGCTTCAGGGG
ACATTGTTCTCCACAGAAGAGGAGCCCTCTAAGCTTCTGGTACAGAACTACCGAGCCCTTCAG
CCTCTCAATCAGCGCATGGTCTTTGCTTCTTTTCATCCAAGCAGGATCCTCGTATACCACAGGT
GAAATGCTGAGTCTAGGTGTAGGAATCTTGGTTGGCTGTCTCTGCCTTCTCCTGGCTGTTTAT
TTCATTGCTAGAAAGATTTCGGAAGAAGAGGCTGGAAAACCGAAAGAGTGTGGTCTTCACCTCA
GCACAAGCCACGACTGAGGCATAAATTCTTCTCAGATACCATGGATGTGGATGACTTCCCTT
CATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTGGCCAGAAACACTGTAGGAGTAG
TAAGCAGATGTCCTCCTTCCCCTGGACATCTCTTAGAGAGGAATGGACCCAGGCTGTCATTCC
AGGAAGAACTGCAGAGCCTTCAGCCTCTCCAAACATGTAGGAGGAAATGAGGAAATCGCTGTG
TTGTTAATGCAGAGANCAAACCTCTGTTTAGTTGCAGGGGAAGTTTGGGATATACCCCAAAGTC
CTCTACCCCTCACTTTTTATGGCCCTTTCCCTAGATATACTGCGGGATCTCTCCTTAGGATAA
AGAGTTGCTGTTGAAGTTGTATATTTTTGATCAATATATTTGGAAATTAAAGTTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPA SYPECGNNAQSPIDIQTDSVTFDPDLPA
LQPHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQIN
SEATFAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQ
KTSVPPFNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTE
EEPSKLLVQNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKI
RKKRLENRKS VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCTGTG
TCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA**ATG**TCC
TACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTCA
TCACAAATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTG
ACAGCTCCAGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATAC
TCCAATCTGAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGT
GTGACCAACCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTG
GAGTCCTTCGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACT
TTGAAAGATCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCT
ATTACCGTGTTTCTTTTTTCTGTGATGGGCTATTCATCTACCGATATATCCACGTGGCAAA
GAGAAACACCCAGCAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTG
CCTGCTGAAAAAATCGTGATTAACCTTTATCACCCCTCAATATCTCGGATGATTCTAAAATTTCT
CATCAGGATATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCC
AGCGGGAACCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCAT
TTGATGGAAATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCTCTCACCCAGCAAGAG
TCCCTCAGCAGAACAATACCCCCGGATAAAACAGTCATTGAATATGAATATGATGTCAGAACC
ACTGACATTTGTGCGGGGCTGAAGAGCAGGAGCTCAGTTTGCAGGAGGAGGTGTCCACACAA
GGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCCGAAACGTTACAGTACTCA
TACACCCCTCAGCTCCAAGACTTAGACCCCTGGCGCAGGAGCACACAGACTCGGAGGAGGGG
CCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGATCCCCAACTGGCAGGCTGTGTATT
CCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAGCCTTCTGAGGGGGATGGGCTC
GGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCTCCAGACAGGCCACCAGGAGAA
AATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTATATGTGCAGATGGAAAAC**TGA**
TGCCAACACTTCCTTTTGCCTTTTGTTCCTGTGCAAACAAGTGAGTCACCCCTTTGATCCCA
GCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGAGTGTCTGTGAGAATTACTT
ATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTTCATGCATGTAGGTCTCTTAACAA
TGATGGTGGGCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAAAGCAGTCAATA
AATGTTTGCCAGACTGGGTGCAGAATTTATTTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQ
IYSNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCA
RTLKDQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRIHVGKEKHPANLILYGNEDKRF
FVPAEKIVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYA
SHLMEIFCDSEENTEGTSLTQQESLSRTIPDKTIVIEYDYDVRTTDICAGPEEQELSLQEEVS
TQGTLLSQAALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRL
CIPSLSSFDQDSEGCEPSEGDLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGAC
ACCTGGGAAGATGGCCGGCCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACCTT
GATCCAAGCCACCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCCCAAAAGTCATCAAAGAAAA
GCTGACACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAG
TGCCATGCGGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCT
GAAGCACATCATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTC
GGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCC
CCTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGA
CACCAGTGCAAGTGGCCCCACCCGCCTGGTCCTCAGTGACTGTGCCACCAGCCATGGGAGCCT
GCGCATCCAAGTCTGTATAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTCATGAA
CCTCCTAGTGCCATCCCTGCCCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTC
CTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGA
CCGTCTGGAGTTTGACCTTCTGTATCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGG
GGCCAAGTTGTTGGACTCACAGGGAAAGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCT
GACAATGCCCACCCTGGACAACATCCCGTTGAGCCTCATCGTGAGTCAGGACGTGGTGAAAGC
TGCAGTGGCTGCTGTGCTCTCTCCAGAAGAATTCATGGTCCTGTTGGACTCTGTGCTTCCTGA
GAGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGAAAAGGCTGCAGATAAGCTGGG
ATCTACCCAGATCGTGAAGATCCTAACTCAGGACACTCCCGAGTTTTTTATAGACCAAGGCCA
TGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTTTCCCTCCAGTGAAGCCCTCCGCCCTTT
GTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGTTTTACACCAAAGGTGACCAACTTAT
ACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTGATGAACTCTGGGATTGGCTGGTT
CCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCACTCCATCCTGCTGCCGAACCA
GAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGGCCTTGGGATTCGAGGCAGC
TGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCCTTGTGGAAACCCAGCTC
TCCTGTCTCCCAGTGAAGACTTGATGGCAGCCATCAGGGAAGGCTGGGTCCCAGCTGGGAGT
ATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTGCCTGTGAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIEKLTQELKDHNATSILQQLPLLSAMR
EKPAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVK
TIVEFHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLV
PSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITQLYLGA
LDSQGKVTKWFNNSAASLTMPITLDNIPFSLIVSQDVVKAABAVALSPEEFMVLLDSVLPESA
HLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTL
GIEASSEAQFYTKGDQLILNLNLISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGK
LRSGVPVSLVKALGFEEAESSLTkdALVLTpaslWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGA
GCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCC**ATGGC**
CTCTCTTGGCCTCCAACCTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGT
TGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTGGTGCCAGCATTGTGACAGCAGT
TGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGA
CATCTATAGCACCCCTTCTGGGCCTGCCCCGCTGACATCCAGGCTGCCCAGGCCATGATGGTGAC
ATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTT
CTGCCAGGAATCCCGAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGG
AGGCCTCCTGGGATTCATTCCTGTTGCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTC
ACCACTGGTGCCTGACAGCATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCT
TTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTTCTGCTCATCCCAGAGAAATCG
CTCCAACCTACTACGATGCCTACCAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGG
TCAACCTCCCAAAGTCAAGAGTGAGTTCAATTCCTACAGCCTGACAGGGTATGTG**TGA**AGAAC
CAGGGGCCAGAGCTGGGGGGTGGCTGGGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCA
CAGGTGAGGGACACTACCACTGGATCGTGTGAGAAGGTGCTGCTGAGGATAGACTGACTTTGG
CCATTGGATTGAGCAAAGGCAGAAATGGGGGCTAGTGTAACAGCATGCAGGTTGAATTGCCAA
GGATGCTCGCCATGCCAGCCTTTCTGTTTTCTCACCTTGCTGCTCCCCTGCCCTAAGTCCCC
AACCTCAACTTGAAACCCCATTCCTTAAGCCAGGACTCAGAGGATCCCTTTGCCCTCTGGT
TTACCTGGGACTCCATCCCCAAACCCACTAATCACATCCCCTGACTGACCCTCTGTGATCAA
AGACCCTCTCTCTGGCTGAGGTGGCTCTTAGCTCATTGCTGGGGATGGGAAGGAGAAGCAGT
GGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCCTCCAAAGAACTGATTGGCCC
TGGAACCTCCATCCCCTCTTGTATGACTCCACAGTGTCCAGACTAATTTGTGCATGAACTG
AAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATGCAGGATGGGAGGACAGGAA
GGCAGCCTGGGACATTTAAAAAATA

FIGURE 80

MASLGLQLVGYYLGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQ
CDIYSTLLGLPADIQAAQAMMVTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFI
LGGLLGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQR
NRSNYYDAYQAQPLATRSSPRPGQPPKVKSEFNSSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTT
TCCCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCCTT
TCCTGCCCACCGCTGCTTCCTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGT
CTGTGGGTTGATCTGTGGCCCTGTGCCTCCGTGTCCTTTTCGTCTCCCTTCCTCCCGACTCC
GCTCCCGGACCAGCGGCCTGACCCTGGGGAAAGGATGGTTCCCGAGGTGAGGGTCCTCTCCTC
CTTGCTGGGACTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCAGACAT
GTTCTGCCTTTTCCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCC
ACAAGGCCTGATGTACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCG
CCTCCACTGTCCGCCTGTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAA
GTGTGTGGAACCTCACACTCCCTCTGGACTCCGGGCCCCACCAAAGTCCTGCCAGCACAACGG
GACCATGTACCAACACGGAGAGATCTTCAGTGCCCATGAGCTGTTCCCCTCCCGCCTGCCCAA
CCAGTGTGTCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCGGCCTCACAACTGCCCCGA
ACCAGGCTGCCCAGCACCCCTCCCACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGC
AAGTGAGCAATCGGATGAAGAGGACAGTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGA
TCCATGTTCCAGTGATGCTGGGAGAAAGAGAGGCCCGGGCACCCAGCCCCACTGGCCTCAG
CGCCCCCTCTGAGCTTCATCCCTCGCCACTTCAGACCCAAGGGAGCAGGCAGCACAACCTGTCAA
GATCGTCCTGAAGGAGAAACATAAGAAAGCCTGTGTGCATGGCGGGAAGACGTACTCCCACGG
GGAGGTGTGGCACCCGGCCTTCCTGTGCCTTCGGCCCTTGCCCTGCATCCTATGCACCTGTGA
GGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCCCACCGAGTACCCCTGCCGTACCCCGA
GAAAGTGGCTGGGAAGTGCTGCAAGATTTGCCCAGAGGACAAAGCAGACCCTGGCCACAGTGA
GATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTCCTCGTCCACACATCGGTATCCCC
AAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGCCTCGGACTTGGTGGAGATCTA
CCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAGGTGAAGTACCTGGCCCAAG
GCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAGGAAGCAAGACTTCCAGA
AAGAGGCACAGCACTTCGACTGCTCGCTGGCCCCCACGAAGGTCCTGGAACGTCTTCCTAG
CCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAGACATAACAAAGAC
CTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAATAAGAAGTTGC
ATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPD MFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTC
SEGAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSA
HELFPSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQ
SLHGVRHPQDPCSSDAGRKRGPPTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKAC
VHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPT EYPCRHPK VAGKCKICP
EDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETE
AQRGEVPGPRPHSQNLPLDSDQESQEARLPERGTALPTARWP PRRSLERLPSPDPGAEGHGQS
RQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAGA
GCCTCTCCGTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCATCC
GTCACCTCTCCTGTCATCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGGCTC
TCATGCTCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGC
CAGACAAGCCTGTCCAGGCCTTGGTGGGGGAGGACGCAGCATTCTCCTGTTTCCTGTCTCCTA
AGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGGCCAGTTCTCTAGCGTGGTCCACC
TCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATCAAGGCAGGACAAAAC
TGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACTGTGTTGG
ATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAAGGCCATCTGGGAGC
TACAGGTGTCAGCACTGGGCTCAGTTCTCTCATTTCATCACGGGATATGTTGATAGAGACA
TCCAGCTACTCTGTCAGTCTCCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTGGAAGGTCCAC
AAGGACAGGATTTGTCCACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGG
AGATCTCTCTGACCGTCCAAGAGAACGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATC
TGAGCCGAGAGGTGGAATCCAGGGTACAGATAGGAGATACCTTTTTTCGAGCCTATATCGTGGC
ACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTATTTTTTGGCATTGTTGGACTGA
AGATTTTCTTCTCCAAATTCCAGTGGAATAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACG
GACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGG
CTCACCCGAAGCTCTGCGTTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCCAGGAGG
TGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAG
GGAAACATTACTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCGTGGGAGTGTGCCGGG
ATGATGTGGACAGGAGGAAGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCTCA
GACTGAATGGAGAACATTTGTATTTACATTAAATCCCCGTTTTATCAGCGTCTTCCCCAGGA
CCCCACCTACAAAATAGGGGTCTTCCTGGACTATGAGTGTGGGACCATCTCCTTCTTCAACA
TAAATGACCAGTCCCTTATTTATACCCTGACATGTGCGGTTTGAAGGCTTATTGAGGCCCTACA
TTGAGTATCCGTCTTATAATGAGCAAAATGGAAGTCCCATAGTCATCTGCCAGTCACCCAGG
AATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCCAGAGACAAGCAACAGTGAGT
CCTCCTCACAGGCAACCACGCCCTTCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCA
CATTCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGGCCAA
GGTGGCTTCCAGATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCCT
GAGCTGGGAGGGAAGAAGGCTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTAAGTGA
TCTTGAAATACCACCTCTCAGGTGAAGAACCGTCAGGAATTCCCATCTCACAGGCTGTGGTGT
AGATTAAGTAGACAAGGAATGTGAATAATGCTTAGATCTTATTGATGACAGAGTGTATCCTAA
TGGTTTGTTCATTATATTACACTTTCAGTAAAAAA

FIGURE 84

MALMLSLVLSLLKLGSQWQVFGPDKPVQALVGEDAAFCFLSPKTNAEAMEVRFFRGQFSSV
VHLYRDGKDQPFMQMPQYQGR TKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAI
WELQVSALGSVPLISITGYVDRDIQLLCQSSGWFFRPTAKWKGPQGQDLSTDRTNRDMHGLF
DVEISLTVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIV
GLKIFFSKFQWKIQAELDWRRKHGQAELRDARKHAVEVTLPETAHPKLCVSDLKTVTHR KAP
QEVPHSEKRFRTRKSVVASQS FQAGKHYWEVDGGHNKRWRVGVC RDDVD R RKEYVTLS PDHGYW
VLR LN GEHLYFTLNPRFISVFPRTPPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLR
PYIEYPSYNEQNGTPIVICPVTQESEKEASWQRASAI PETSNSESSSQATT PFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCCAGACATGCTGCTGCTGCTGCTGCCC
CTGCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCC
GTGACGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGC
TGGATTTACCCTGGCCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAG
GATGCTCCAGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTTC
CACCTCCTTGGGGACCCACATAACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGT
GATGCGGGGAGATACTTCTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCAC
CGGCTCTCTGTGAATGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCAGGCACCCTG
GAGTCCGGCTGCCCCCAGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCC
CCTATGATCTCCTGGATAGGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCG
GTGCTCACCCCTCATCCACAGCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTC
CCTGGGGCCAGCGTGACCACGAACAAGACCGTCCATCTCAACGTGTCCTACCCGCCTCAGAAC
TTGACCATGACTGTCTTCCAAGGAGACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCT
CTGTCACTCCCAGAGGGCCAGTCTCTGCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAAT
CCCCCTGCCAGGCTGAGCCTGAGCTGGAGAGGCCTGACCCTGTGCCCTCACAGCCCTCAAAC
CCGGGGGTGCTGGAGCTGCCTTGGGTGCACCTGAGGGATGCAGCTGAATTCACCTGCAGAGCT
CAGAACCCTCTCGGCTCTCAGCAGGTCTACCTGAACGTCTCCCTGCAGAGCAAAGCCACATCA
GGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAGCCACAGCCCTGGTCTTCCTGTCCTTCTGC
GTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAATCGGCAAGGCCAGCAGCGGGCGTGGGA
GATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTTCAGCCTCTCAGGGGGCCCCTGACTGAA
CCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCAGCTTCTGCCCCTCCTCAGTGGGG
GAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCAGATGGTGAAGCCTTGGGACTCGCGGGGA
CAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAGATTGAGAAACTGCAGAGACT
CACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGAGAAGTCAGAGGCTGATTCTTGTA
AATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATGTGCAGAGTGAAAAGC
ACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCCTCCCTTTTATTTT
TTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSSYP SHGWIYPGPVVHGYWFR
EGANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGS I
KWN YKHHRLSVNV TALTHRPNILIPGTLESGCPQNL TCSVPWACEQGTPPMISWIGTSVSPLD
PSTTRSSVLT LI PQPDHGTSLTCQVT FPGASVTTNKT VHLNVSYPPQNL TMTVFQGDGT VST
VLGNGSSLSLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDA
AEFTCRAQNPLGSQQVYLNVS LQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSA
RPAAGVGDTGIEDANAVRG SASQG PLTEPWAEDSPPDQPPPASARSSVGEGELQYASLSFQMV
KPWDSRGQEATDTEYSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGT
ACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACA**ATGA**ACCAACTCAGCTTCCTGCTG
TTTCTCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGG
ACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGT
GCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGAC
ATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGG
AAGTGCACGGTGGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGG
GACGGCAACTGGGCCAACTACAACACCTTTGGATCTGCAGAGGGCGGCCACGAGCGATGACTAC
AAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAG
TCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTC
CAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGA
AAGTGTTGGACTGACAACGGCCCGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCCAGAAA
ACAGCATCTTATTACTCACCTATGGCCAGCGGGAATTCAGTGCGGGATTTGTTTCAGTTCAGG
GTATTTAATAACGAGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAAC
ACTGAGCATCACTGCATTGGTGGAGGAGGATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGA
GATTTTCTGGTTTTTGATTGGAGTGGATATGGAACCATGTTGGTTACAGCAGCAGCCGTGAG
ATAACTGAGGCAGCTGTGCTTCTATTCTATCGT**TGA**GAGTTTTGTGGGAGGGAACCCAGACCT
CTCCTCCAACCATGAGATCCAAGGATGGAGAACAACCTACCCAGTAGCTAGAATGTTAATG
GCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENG
VIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSA
EAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQK
YPVKYGEGKCWTDNGPVI PVVYDFGDAQKTASYYSYPYGQREFTAGFVQFRVFNNERAAALCA
GMRVTGCNTEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSRSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAG
ACCGCCGCCCTTGTCCCCGAGGGCC**ATG**GGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCT
GACGCTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACA
GGCCTGCCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGC
CGCGCTCTCTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGAGTCTC
CATGTTCAACAGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCT
GTCCTTCTTCATATTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTTGTCTTCTGCAG
TGCCCTTCCAGCTGTCACTGAAATGGCTTTATTCGTCACCGTCTTTGGGCTGAAAAAGAAACC
CTTC**TGA**TTACCTTCATGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGT
ATTCCTGGAAGAAGGAAGGCATAGGCTTCGGTTTTCCCCTCGGAAACTGCTTCTGCTGGAGGA
TATGTGTTGGAATAATTACGTCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAA
TAAAATATGTTTTGTAGTAACATTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAA
AAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGL
FAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEM
ALFVTVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

12-28
51-66
107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAG**ATG**CAACTG
 ACTCGCTGCTGCTTCGTGTTCCCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGAT
 GATGGTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGG
 GTGCCTCGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTA
 GGGCTGCTGGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCCCGAAC
 CACAGCCCCCACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAAC
 ATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACC
 TTCAGCGTCCACTTCCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCC
 CCCAGTAAAGCTGTAGAGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAA
 ATCTTCAACTGCCGGATGGAGTGGGAGAAGGTAGAACGGGGCCGCCGGACCTCGCTTTGCACC
 CACGACCCAGCCAAGATCTGCTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCC
 CAGCCCTTCAAAGTCGTCTGTGTCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAG
 AAGGTGTGCCCAGATTACAACCTACCATAGTGATACCCCCTACTACCCATCTGGG**TGA**CCCCGGG
 GCAGGCCACAGAGGCCAGGCCAGGGCTGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGG
 ACACCGGGCAGGGAAGGGGTGGGCCTCAGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAA
 GTGGGGCCAGGGCCAAGTCTCAAGTGGCAGAGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAA
 GCTGTGGAGTGACTAGATCACAGGAGCACTGGAGGAGGAGTGGGCTCTCTGTGCAGCCTCACA
 GGGCTTTGCCACGGAGCCACAGAGAGATGCTGGGTCCCCGAGGCCTGTGGGCAGGCCGATCAG
 TGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGCCCTTGGTTCTTGCCATCCTGAGGAAAG
 ATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACAGCCTGTCAACTTAGGATGGATGGCT
 GAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGGGCCAGAGGAGCTCTCCAGCCCTG
 CCTAGTGGGCGCCCTGAGCCCTTGTCTGTGTGCTGAGCATGGCATGAGGCTGAAGTGGCAACC
 CTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGCCAGGCCACCCCTTTCCAAA
 ATTCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATGGCACACCCATCCTTAAG
 CTAAGACAGGACGATTGTGGTCCTCCCACACTAAGGCCACAGCCCATCCGCGTGCTGTGTGTC
 CCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCGGAGAGGGGTCCCT
 CAACAGTCAGCCTCACCTGTCAGACCGGGGTCTCCCGGATCTGGATGGCGCCGCCCTCTCAG
 CAGCGGGCACGGGTGGGGCGGGGCCGGGCCGAGAGCATGTGCTGGATCTGTTCTGTGTGTCT
 GTCTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTTTGTGTGA
 AGAATCGTGTTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANS
TLLGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHG
NGTFSVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTS
LCTHDPAKICSRDHAQSSATWSCSQPFKVVVCVYIAFYSTDYRLVQKVCPCDYNHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCC**ATG**ACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCG
CCCTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTT
TCTTCTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTG
ACAACAAAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATA
TCCAAGAAATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGA
GTATAAACCCAGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCT
TTGGAATCATGAGTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCA
CAGTGGGCATTTCATGGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTCATGACGCTGGTCA
TTATCTTGCTGCATGTATTCTGGGGCATTGTATTTTTTTGATGGCTGTGAGAAGAAAAAGTGGG
GCATCCTCCTTATCGTTCTCCTGACCCACCTGCTGGTGTCAGCCCAGACCTTCATAAGTTCTT
ATTATGGAATAAACCTGGCGTCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCT
TAGCTGCGGGAGGCAGCTGCCGAAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTC
TTCTTTACAACCAGCGCTCCAGAT**TAA**CCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACA
TCTTTAGAGGAAGCACAACTGTGCCTTTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAA
GAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNK
DGPTQKYLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGI
MSGVFSFVNTLSDSLPGPTVGIHGDSPOFFLYSAFMTLVIILLHVFVGIVFFDGCCKKKGIL
LIVLLTHLLVSAQTFISSYYGINLASAFIILVLMGTWAFILAAGGSCRSCLKCLLCQDKNFLLY
NQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

[illegible]

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTLISSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190

FIGURE 97

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCT
 CTCTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCCTGAC
CATGGTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGCCCA
 GCCTGCAGAGCTGTCTGTGGAAGTTCCAGAAAATATGGTGGAATTTCCCTTTATACCTGAC
 CAAGTTGCCGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTCAGGGGACTCAGGCAA
 GGCAACTGAGGGCCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAGGGCCCT
 GGACCGAGAGGAGCAGGCAGGTACCAGCTACAGGTCACCCTGGAGATGCAGGATGGACATGT
 CTTGTGGGGTCCACAGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTT
 CTCTCAAGCCATCTACAGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCCCCTTCCTCTT
 CCTTGAGGCTTCAGACCGGGATGAGCCAGGCACAGCCAACTCGGATCTTCGATTCCACATCCT
 GAGCCAGGCTCCAGCCCAGCCTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGGCTCT
 GGCCCTCAGCCCCAAGGGGAGCACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTT
 GGTACAGGTCAAGGACATGGGTGACCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAGT
 CTCCATCATAGAGAGCACCTGGGTGTCCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGT
 CCTATACCCGCACCACATGGCCCAGGTACACTGGAGTGGGGGTGATGTGCACTATCACCTGGA
 GAGCCATCCCCCGGGACCCCTTTGAAGTGAATGCAGAGGGAAACCTCTACGTGACCAGAGAGCT
 GGACAGAGAAGCCCAGGCTGAGTACCTGCTCCAGGTGCGGGCTCAGAATTCCCATGGCGAGGA
 CTATGCGGGCCCTCTGGAGCTGCACGTGCTGGTGATGGATGAGAATGACAACGTGCCTATCTG
 CCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCTCAGTCCACCAGGTACTGAAGTGACTAG
 ACTGTCAGCAGAGGATGCAGATGCCCCCGGCTCCCCCAATTCCACGTTGTGTATCAGCTCCT
 GAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTCCAGGTGGACCCCACTTCAGGCAG
 TGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGCAACATCCTGCTTCTGGTGCTGGCCAT
 GGACCTGGCAGGCGCAGAGGGTGGCTTCAAGCAGCAGCTGTGAAGTCGAAGTCGCAGTCACAGA
 TATCAATGATCACGCCCCTGAGTTCATCACTTCCAGATTGGGCCTATAAGCCTCCCTGAGGA
 TGTGGAGCCCGGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGACCTCGAGCCCGCCTT
 CCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTGGCCTGGATTGGGA
 GCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAGGCAGCTCCAAG
 TCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCCAGGCCCTGG
 AGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGATGCCACCCCCAAGTTGGACCAGGA
 GAGCTACGAGGCCAGTGTCCTCATCAGTGCCCCAGCCGGCTCTTTCCTGCTGACCATCCAGCC
 CTCCGACCCCATCAGCCGAACCCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCTCTG
 CATTGAGAAATTCTCCGGGGAGGTGCACACCGCCCAGTCCCTGCAGGGCGCCAGCCTGGGGA
 CACCTACACGGTGCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCCCTGTGCCCTCCCA
 ATACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGA
 TCTGGCCAGTGGGCACGGTCCCTACAGCTTACCCTTGGTCCCAACCCACGGTGCAACGGGA
 TTGGCGCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGA
 GCCACGTGAACACATAATCCCCGTGGTGGTCAGCCACAATGCCAGATGTGGCAGCTCCTGGT
 TCGAGTGATCGTGTGTCGCTGCAACGTGGAGGGGCAGTGATGCGCAAGGTGGGCCGATGAA
 GGGCATGCCACGAAGCTGTCGGCAGTGGGCATCCTTGTAGGCACCCTGGTAGCAATAGGAAT
 CTTCTCATCCTCATTTTACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGC
 AGACAGCGTGCCCTGAAGGCGACTGT**TGA**ATGGCCAGGCAGCTCTAGCTGGGAGCTTGGC
 CTCTGGCTCCATCTGAGTCCCCTGGGAGAGAGCCAGCACCCAAGATCCAGCAGGGGACAGGA
 CAGAGTAGAAGCCCTCCATCTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCT
 GCAGAGCCTGGACACCAACTTTATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTTTG
 CCAATAATAAAGCCCCAGAGAACTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPaelSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGK
ATEGPFAMDPDSGFLLVTRALDREEQAeyQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHF
SQAIYRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSDFQLEPRLGAL
ALSPKGSTSLDHALERTYQLLVQVKMDGDAQSGHQATATVEVSIESTWVSLEPIHLAENLKV
LYPHHMAQVHWSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAeyLLQVRAQNSHGED
YAAPLELHVLVMDENDNVPICPPRDPTVSIPELSPPGTEVTRLsAEDADAPGSPNSHVYQLL
SPEPEDGVEGRAFQVDPTSGSVTLGVLPRLAGQNILLVLAMDLAGAEGGSSTCEVEVAVTD
INDHAPEFITSQIGPISLPEDVEPGTLVAMLTADADLEPAFRLMDFAIERGDTEGTFGLDWE
PDSGHVRLRLCKNLSYEAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQE
SYEASVPISAPAGSFLLTIQSPDPIsRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGD
TYTVLVEAQDTALTAPVPSQYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRD
WRLQTLNGSHAYLTALHWWVEPREHIIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMK
GMPTKLSAVGILVGTlVAIGIFLILIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCC
TGAGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCC
TTTATCTCTTCACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCT
TGGTACATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAA
AGGAAATGTTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTTCCAATGA
GACTAGCACCTCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAA
CTCTGGGTCCAGTGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGAC
CTCCAATGGGGTCAGCATAGTCACCAACTCTGAGTTCCATAACCTCCAGTGGGATCAGCAC
AGCCACCAACTCTGAGTTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTC
CAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGG
GGCCAGCACAGTCACCAACTCTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAA
CTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACT
CTCCAGTGGGGCCAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCAC
AGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTC
CAGCACAGTGTCCAGTAGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGAACGACCTCCAATGGGGCTGGCACAGCCACCAA
CTCTGAGTCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGACTCCAGCACAGT
GTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCAGCAC
AGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCTAGCACAGCCACCAACTCTGACTC
CAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTGG
GATCAGCACAGTCACCAATTCTGAGTCCAGCACACCTCCAGTGGGGCCAAACACAGCCACCAA
CTCTGAGTCCAGTACGACCTCCAGTGGGGCCAAACACAGCCACCAACTCTGAGTCCAGCACAGT
GTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGTCCAGCAC
AGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGACTC
CAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAGCACAGTGTCCAGTGG
GATCAGCACAGTCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCAAACACAGCCACCAA
CTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGAATGCACACAAC
TTCCCATAGTGCATCTACTGCAGTGAAGTGAAGCAAAGCCTGGTGGGTCCCTGGTGCCGTGGGA
AATCTTCCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGCTCTTCTT
CTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCTCATGG
CCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCACAGGCCAGGTG
GAGTCCTAACTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGAA
CAGCGGGCCCTTGAAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGC
ACCCAAGACCTGGTTTCCTTTTCATTCATCCCAGGAGACCCCTCCAGCTTTGTTTGAGATCCT
GAAAATCTTGAAGAAGGTATTCTTCACCTTTCTTGCCCTTACCAGACACTGGAAAGAGAATAC
TATATTGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCT
GTGCTTGCCCCGGGGTGGGTATCTAGCTCTGAGATGAACTCAGTTATAGGAGAAAACCTCCAT
GCTGGACTCCATCTGGCATTCAAAATCTCCACAGTAAAATCCAAAGACCTCAAAAAAAAAAAAA
AAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATIS
GSSVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESS
TPSSGASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTS
SGASTATNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGA
GTATNSESSTTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTA
TNDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNS
ESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESS
TVSSGISTVTNSESSTTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGS
LVPWEIFLITLVSVVAAVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAP
HRPRWSPNWFWRPVS SIAMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCGCGACGGAGGTTGTGACCCC
TACGGAGCCCCAGCTTGCCACGCACCCCCTCGGCGTCGCGCGGCGTGCCCTGCTTGTGACA
GGTGGGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCA
ACAAATGGATGATGTGATAT**ATG**CATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCAT
GGTCAATTAACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCT
TTGGAATCATGGTGTGATGGAAAGGGATTTACTTTTATACTGACTCTGTTTTGGGGAAGCTTTT
TTGGAAGCATTTTCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATC
GCTGGATCAACAACCGCCTTGTGGCAACATGGCTCACCCCTACCTGTGGCATTATTGGAGACCA
TGTTTGGTGTAAGAGTGATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCATTA
TCATGAACCATCGGACAAGAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCT
ACCTCAGATTGGAGAAAATTTGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTTGGTTGGG
CCATGCAGGCTGCTGCCTATATCTTCATTCATAGGAAATGGAAGGATGACAAGAGCCATTTGCG
AAGACATGATTGATTACTTTTGTGATATTCACGAACCACTTCAACTCCTCATATTTCCAGAAG
GGACTGATCTCACAGAAAACAGCAAGTCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTC
AGAAATATGAATATGTTTTACATCCAAGAACTACAGGCTTTACTTTTGTGGTAGACCGTCTAA
GAGAAGGTAAGAACCTTGATGCTGTCCATGATATCACTGTGGCGTATCCTCACAACATTCCTC
AATCAGAGAAGCACCTCCTCCAAGGAGACTTTCCCAGGGAAATCCACTTTCACGTCCACCGGT
ATCCAATAGACACCCTCCCCACATCCAAGGAGGACCTTCAACTCTGGTGCCACAAACGGTGGG
AAGAGAAAGAAGAGAGGCTGCGTTCTCTATCAAGGGGAGAAGAATTTTTATTTTACCGGAC
AGAGTGTCATTCCACCTTGCAAGTCTGAACTCAGGGTCCTTGTGGTCAAATTGCTCTCTATAC
TGTATTGGACCCTGTTTCAGCCCTGCAATGTGCCTACTCATATATTTGTACAGTCTTGTTAAGT
GGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGAGAATATTTGGTGGACTGGAGA
TCATAGAACTTGCATGTTACCGACTTTTACACAAACAGCCACATTTAAATTCAAAGAAAAATG
AG**TAA**GATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTGGAATGTTCTAAACCTT
TCTAAGCTCAGATGCATTTTTGCATGACTATGTGCAATATTTCTTACTGCCATCATTATTTGT
TAAAGATATTTTGCACCTAATTTTGTGGGAAAAATATTGCTACAATTTTTTTTAAATCTCTGAA
TGTAATTTTCGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTTGGGCCAGAATA
TTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFSGIMVSWKGIYFILTLFWGSFFGSI FML
 SPFLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDADFVPGERSVIIMNHRTR
 MDWMFLWNCLMRYSYLRLEKICLKASLKGVPFGFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYF
 CDIHEPLQLLIIFPEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLD
 AVHDITVAYPHNIPQSEKHL LQGDFPREIHFHVHRYPIDTLPTS KEDLQLWCHKRWEEKEERL
 RSFYQGEKNFYFTGQSVIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITI
 VIFVLQERIFGGLEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

102/168

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGA
 CCTCCAAATCATCCATCCACCCCTGCTGTCTATCTGTTTTTCATAGTGTGAGATCAACCCACAGG
 AATATCC**ATG**GCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTG
 GCAAGTCACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCCTG
 CTCCCTCTTTCTGAGACCAGTGCAGAGGCTATGGAAGTGCAGTTCTTCAGGAATCAGTTCCA
 TGCTGTGGTCCACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCG
 AGGGAGAAGTGAAGTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGCTAAAAA
 CATCACTCCCTCGGACATCGGCCTGTATGGGTGCTGGTTTCAGTTCCAGATTTACGATGAGGA
 GGCCACCTGGGAGCTGCGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCATCGTGGGATA
 TGTTGACGGAGGTATCCAGTTACTCTGCCTGTCTCAGGCTGGTTCCCCCAGCCACAGCCAA
 GTGGAAAGGTCCACAAGGACAGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAG
 CCTGTATGATGTGGAGATCTCCATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCAT
 CCACCTTGCTGAGCAGAGTCACTGAGGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCA
 GCCCTCACCTTGGCGCCTGGCTTCTATTTTACTCGGGTTACTCTGTGGTGGCCTGTGTGGTGT
 TGTCATGGGGATGATAATTGTTTTCTTCAAATCCAAAGGGAAAATCCAGGCGGAACTGGACTG
 GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGACTCT
 GGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTGAACCCATAGAAA
 AGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCA
 GGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGACGTGGGACAAAATGTAGGGTGGTATGT
 GGGAGTGTGTGCGGATGACGTAGACAGGGGGAAGAACAATGTGACTTTGTCTCCCAACAATGG
 GTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTACATTCAATCCCCATTTTATCAG
 CCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCTCGGACTATGAGGGTGGGACCAT
 CTCTTTCTTCAATACAAATGACCAGTCCCTTATTTATACCCTGCTGACATGTCAGTTTGAAGG
 CTTGTTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAGGGGACTCCCATATTTCAT
 ATGTCCAGTGTCTGCGGAT**AG**ACAGAGAGAAGACCCTGCTTAAAGGGCCCCACACCACAGACC
 CAGACACAGCCAAGGGAGAGTGTCTCCCGACAGGTGGCCCCAGCTTCCTCTCCGGAGCCTGCGC
 ACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCCTGAGCCCTGCA
 GCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAGTCAGAAGCC
 ATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAACCTCCATC
 CAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTCACGGACAGT
 GATTCTGCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTTTGA
 GGGCACAGTGTGTTGCTAATGATGTGTTTTTATATTATACATTTTCCCACCATAAACTCTGTTT
 GCTTATTCCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACAC
 CTGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTCTACTGATTCTATAAGCCAG
 CATTACCTGATACCAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCA
 TATCCCTCATTAACACAGACACAAAAATTCTAAATAAAATTTTAACAAATTAACTAAACAAT
 ATATTTAAAGATGATATATACTACTCAGTGTGGTTTGTCCCACAAATGCAGAGTTGGTTTAA
 TATTTAAATATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAGAAAACCATAAAAAA
 AAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAV
VHLYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEAT
WELRVAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLY
DVEISIIVQENAGSILCSIHLEQSHVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVM
GMIIVFFKSKGKIQAELDWRRKHGQAELRDARKHAVEVTLPETAHPKLCVSDLKTVTHRKAP
QEVPHSEKRFRTRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDVRGKNNVTLSPNNGYW
VLRLTTEHLYFTFNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLL
RPYIQHAMYDEEKGTPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCA**ATG**ATGTATCGGCCAGATGTGGTGAGGGCTAG
GAAAAGAGTTTGTGGGAACCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCCT
GGCAGTGTGCATTGGACTCACTGTTTATTATGTGAGATATAATCAAAGAAGACCTACAATTA
CTATAGCACATTGTCAATTTACAAGTACAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAA
CAATTTTACAGAAATGAGCCAGAGACTTGAATCAATGGTGAAAAATGCATTTTATAAATCTCC
ATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTCAGTCAACAGAAGCATGGAGTGTT
GGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGAAACTGTAGATAAAAT
TGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAGATCCTCA
CTCAGTTAAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGG
AACACGAAGAAGTAAACTCTAGGTGAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGA
GGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTT
AATTAATGCCACATGGCTTGTGAGTGCTGCTCACTGTTTTACAACATATAAGAACCCTGCCAG
ATGGACTGCTTCCTTTGGAGTAACAATAAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAAT
AATTGTCCATGAAAAATACAAACACCCATCACATGACTATGATATTTCTCTTGACAGAGCTTTC
TAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCTCCCTGATGCATCCTATGAGTT
TCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAATGATGGTTACAGTCA
AAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACTTGCAATGAACCTCAAGC
TTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGC
ATGCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGC
TGGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGTTTATACTAGAGT
TACGGCCTTGCGGGACTGGATTACTTCAAAAAGTGGTATC**TAA**GAGACAAAAGCCTCATGGAA
CAGATAACATTTTTTTTTTGTTTTTTGGGTGTGGAGGCCATTTTTAGAGATACAGAATTGGAGA
AGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAACTGTTTGCTTGATGCATGTATTTT
CTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCAGATCAACTCTGTCATCTGTGA
GCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATACAATATTACATTACAGCCTGTA
TTCATTTGTTCTCTAGAAGTTTTGTGAGAATTTTGAAGTTGTTGACATAAATTTGTAATGCATA
TATACAATTTGAAGCACTCCTTTTCTTCAGTTCCTCAGCTCCTCTCATTTTACAGCAAATATCCA
TTTTCAAGGTGCAGAACAAAGGAGTGAAAGAAAAATATAAGAAGAAAAAATCCCCTACATTTTA
TTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGTGGAATATTAGAAATGATCATATTCATTAT
GAAAGGTCAAGCAAAGACAGCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACTAA
GTTAAGGAAGTCCAGAAAGAAGCCAAGATATATCCTTATTTTCATTTCCAAACAACTACTATG
ATAAATGTGAAGAAGATTCTGTTTTTTTGTGACCTATAATAATTATACAACTTCATGCAATG
TACTTGTTCTAAGCAAATTAAGCAAATATTTATTTAACATTGTTACTGAGGATGTCAACATA
TAACAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYSTLSFTTDK
LYAEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQQKHGVLAHMLLICRFH
STEDPETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQS
LRIVGGTEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIK
PSKMKRGLRRIIVHEKYKHPSHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDVMFVTG
FGALKNDGYSONHLRQAQVTLI DATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLV
SSDARDIWYLAGIVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCC
TGCCCCGATGAGCCCCCGCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACC GGCC
CAGCGCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTC
TCCCTTACGGGGGCTCACA**ATG**GCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTAC
GCCCTCAATCTGCTCTTTTGGTTAATGTCCATCAGTGTGTGGCAGTTTCTGCTTGGATGAGG
GACTACCTAAATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTG
ACTTACTTTTCTGTGGTTTCATCCGGTCATGATTGCTGTTTGGCTGTTTCTTATCATTGTGGGG
ATGTTAGGATATTGTGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGT
TTGCTTGTCAATTTTCTGTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTTATG
GTTCCAGTACAATGGTCAGATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCT
AGATATCGGTGGCTTACTCATGCTTGGAAATTTTTTTCAGAGAGAGTTTAAAGTGTGTGGAGTA
GTATATTTCACTGACTGGTTGGAAATGACAGAGATGGACTGGCCCCCAGATTCTCTGTGTGTT
AGAGAATTCCCAGGATGTTCCAAACAGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAG
GGTTGTGGGAAGAAAATGTATTCTTTTGGAGGAACCAAACAACCTGCAGGTGCTGAGGTTT
CTGGGAATCTCCATTGGGGTGACACAAATCCTGGCCATGATTCTCACCATTACTCTGCTCTGG
GCTCTGTATTATGATAGAAGGGAGCCTGGGACAGACCAAATGATGTCCTTGAAGAATGACAAC
TCTCAGCACCTGTCATGTCCCTCAGTAGAACTGTTGAAACCAAGCCTGTCAAGAATCTTTGAA
CACACATCCATGGCAAACAGCTTTAATACACACTTTGAGATGGAGGAGTTA**TAAAAA**GAAATG
TCACAGAAGAAAACCAAACTTGTTTTTATTGGACTTGTGAATTTTTGAGTACATACTATTCTA
TTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAATAACACCTAAGCATATACTATTCTA
TGCTTTAAATGAGGATGGAAAAGTTTCATGTGCATAAGTCACCACCTGGACAATAATTGATGC
CCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCCTGTGTATGACTTTTACTGAAC
ACAGTTATGTTTTGAGGCAGCATGGTTTTGATTAGCATTTCGCATCCATGCAAACGAGTCACA
TATGGTGGGACTGGAGCCATAGTAAAGGTTGATTTACTTCTACCAACTAGTATATAAAGTACT
AATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTACTCAGCGATCTATTC
TTCTGATGCTAAATAAATTATATATCAGAAAACCTTTCAATATTGGTGACTACCTAAATGTGAT
TTTTGCTGGTTACTAAAATATTCTTACCACTTAAAAGAGCAAGCTAACACATTGTCTTAAGCT
GATCAGGGATTTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTTCAGTCGATTTTCAGTTCT
GATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATTATTTTAA
GCCTTTCCTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAACTGTTA
TTTAAATACTTAACCACTAATTTTGAAAATTACCACTGTGATACATAGGAATCATTATTCAGA
ATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTTCAGAA
AGGACTTGTATGCTGTTTTTCTCCCAAATGAAGACTCTTTTGGACACTAAACACTTTTTAAAA
AGCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTCTACAGA
AAATAGTGTTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAAACATGTGACAATTTAG
AGATTCTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTT
TTTACAAGAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTAT
TTTGTTTTATTTCTCAGAATATGGAAAGAAAATTAAATGTGTCAATAAATATTTTCTAGAGAG
TAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVV
HPVMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWS
DMVTLKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREFPGC
SKQAHQEDLSPLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDR
REPGTDQMMSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTTCCTCTGATAAAGCCC
 CTACCACTGCTGATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAGTGCTTGA
 AAGAGAAGGGGACAAAGGAACACCAGTATTAAGAGGATTTTCCAGTGTTTCTGGCAGTTGGTC
 CAGAAGG**ATG**CCTCCATTCTGCTTCTCACCTGCCTCTTCATCACAGGCACCTCCGTGTCACC
 CGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGCCCTGGAGGAACACTGACCA
 CCAGTTGGATGAGTCTCAAGGTCCCTCTATGTGACAACCATGTGAATGGGGAGTGGTACCA
 CTTACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAAC
 CCACGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAAGGCGACGGCATTGTGCAACGCCA
 GGCTTGTGCCAGCTTCAATGGGAACCTGCTGTCTCTGGAACACCACGGTGGGAAGTCAAGGCTTG
 CCCTGGAGGCTACTATGTGTATCGTGTGACCAAGCCAGCGTCTGCTTCCACGTCTACTGTGG
 TCATTTTTATGACATCTGCGACGAGGACTGCCATGGCAGCTGCTCAGATACCAGCGAGTGCAC
 ATGCGCTCCAGGAACCTGTGCTAGGCCCTGACAGGCAGACATGCTTTGATGAAAATGAATGTGA
 GCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTACCGCTGTGAGTG
 TGGGGTTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAA
 TAACAATGGTGGCTGCAGCCACTCTTGCCTTGGATCTGAGAAAGGCTACCAGTGTGAATGTCC
 CCGGGCCCTGGTGTGTCTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGTGCAAAATCAAA
 TGCCATTGAAGTGAACATCCCCAGGGAGCTGGTGGGCTGGAGCTCTTCTCTCAAGACATGTGG
 CTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACATCCTCTTCTCTCAAGACATGTGG
 TACAGTGGTGCATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGACAGGTCTACCCAA
 GCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTGAC
 CTGCGAGTTTCCACGCCTGTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCC
 ACTGGAAATCATGAGCCGAAATCATGGGATCTTCCCATTCACTCTGGAGATCTTCAAGGACAA
 TGAGTTTGAAGAGCCTTACCGGGAAGCTCTGCCACCCCTCAAGCTTCGTGACTCCCTCTACTT
 TGGCATTGAGCCCGTGGTGCACGTGAGCGGCTTGGAAGCTTGGTGGAGAGCTGCTTTGCCAC
 CCCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCATCCGGGATGGCTGTGTTTCAGA
 TGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCAGGTCCCTGTCTT
 CAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTCTTGTCTGTGGAGT
 GTTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGCGAATGCGTCGTGGGGCAGGAGG
 AGAGGACTCAGCCGGTCTACAGGGCCAGACGCTAACAGGCGGCCGATCCGCATCGACTGGGA
 GGACT**TAG**TTTCGTAGCCATACCTCGAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTC
 CCCCCACCGCCCTCTAAGAACATCTGCCAACAGCTGGGTTTCACTTCACTGTGAGTTCAG
 ACTCCAGCACCAACTCACTCTGATTCTGGTCCATTCACTGGGCACAGGTCACAGCACTGCTG
 AACAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAACCTAACTGTCCACCCAGAA
 AGACACTCACCCCATTTCCCTCATTTCTTTCTTACACTTAAATACCTCGTGTATGGTGAATC
 AGACCACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAAACAG
 TTACTGAAATTATGACTTAAATACCCAATGACTCCTTAAATATGTAAATTATAGTTATACCTT
 GAAATTTCAATTCAAATGCAGACTAATTATAGGGAATTTGGAAGTGTATCAATAAACAGTAT
 ATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFT
 GMAGDAMPTFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPG
 GYYVYRLTKPSVCFHVYCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQN
 NGGCSEICVNLKNSYRCECGVGRVLRSDGKTCEVEGCHNNNGGCSHSCLGSEKGYQCECPRG
 LVLSEDNHTCQVPVLCKSNAIEVNI PRELVGGLELFLTNTSCRGVSNNGTHVNILFSLKTCGTV
 VDVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVTCEFPRLYTISEGYVPNLRNSPLE
 IMSRNHGIFPFTLEIFKDNEFEOPYREALPTLKLRLDSLYFGIEPVVHVSGLESLVESCFATPT
 SKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVVGKDHKEVFLHCRVLVCGVLD
 ERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
 522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCT
GCACTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACC
TGTGTGGGGAGGCCCTCCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACC
GGGAGGATCACAGAGCCAGC**ATGT**TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCG
ATGTCAAACCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGGATCCCCA
TCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGTGGTTGTCTCATCAAGGTGATTCT
TGGATAAATACTACTTCCCTCTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTG
ACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCCGAAG
GGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCA
CAGGGAACCTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTA
GGCAGATGGGCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAA
TCACAGAAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGTCTCTCAGGCTCCC
TGGTCTCCCTGCACTGTCTTGCCCTGTGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGG
AGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGGTGAGCATCCAGTACGACAAACAGCACGTCT
GTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCAGCCCACTGCTTCAGGAAACATA
CCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAGCTTCCCATCCCTGG
CTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAAAGACAATGACATCGCCC
TCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCT
TTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGA
ATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGT
GCAATGCAGACGATGCGTACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGG
AAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCCCTGATGTACCAATCTGACCAGT
GGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGGGGGCCCGAGCACCCCAAGGAGTAT
ACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCT**GTA**ATGCT
GCTGCCCCCTTTGCAGTGCTGGGAGCCGCTTCCTTCCTGCCCTGCCACCTGGGGATCCCCAA
AGTCAGACACAGAGCAAGAGTCCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCATTTCTT
GGAGCAGCAAAGGGCCTCAATTCTGTAAAGAGACCCTCGCAGCCCAGAGGCGCCCAGAGGAAG
TCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAGCATCCCAGGGAGAGACACAGCCCACT
GAACAAGGTCTCAGGGGTATTGCTAAGCCAAGAAGGAACTTTCCCACACTACTGAATGGAAGC
AGGCTGTCTTGTAAGGCCAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGTCTGCGCCA
GCCCTGTCCGTCTTACCCATCCCCAAGCCTACTAGAGCAAGAAACCAGTTGTAATATAAAAT
GCACTGCCCTACTGTTGGTATGACTACCGTTACCTACTGTTGTCATTGTTATTACAGCTATGG
CCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAAAAAAAA

FIGURE 112

MLQDPDS DQPLNSLDVKPLRKPRIPMETFRKVGIP III IALLSLASIIIVVVL IKVILDKYYFL
CGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLD SATGNWFSA
CFDNFTEALAETACRQMGYSRAVEIGPDQDL DVVEITENSQELMRNSSGPCLSGSLVSLHCL
ACGKSLKTPRVVGEEASVDSWPWQVSIQYDKQHVC GGSILDPHWVLTAAHCFRKHTDVFNWK
VRAGSDKLGSFP SLAVAKII IIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTP
ATPLWIIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGV DTC
QGDSSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGAT
TATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTTTTCAGCAACTAAAAAAGCCACAG
GAGTTGAACTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATT
AAAATCTGTTTTTTGTTCTCTTGTAAGTAGCCTTTACCTTCCTAACACAGAGGATCTGTCACT
GTGGCTCTGGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGT
CCCCTCGAAGCCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAAC
TGTCTCACGTCTGGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGG
CTTTCAAGGTGGGCCTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTTTCATAGGCG
ATGGCTCCCCTGCCCAGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGG
GCCGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCCGCCCATCTAACCTTTT
CATGTCCTGCACATCACCTGATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCT
TGAGTGAGCTGTGGCTCAGACCCAGAAGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGA
CTTGCAATTCTCCTGGAACATGAGGGAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACTTG
TGCCAAATTATGGGTGAGAAAAGATGGAGGTGTTGGGTTATCACAAGGCATCGAGTCTCCTGC
ATTCAGTGGACATGTGGGGGAAGGGCTGCCGATGGCGCATGACACACTCGGGACTCACCTCTG
GGGCCATCAGACAGCCGTTTCCGCCCCGATCCACGTACCAGCTGCTGAAGGGCAACTGCAGGC
CGATGCTCTCATCAGCCAGGCAGCAGCCAAAATCTGCGATCACCAGCCAGGGGCAGCCGTCTG
GGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCCCTCCTTCCCTCTGAGAGGCCCTCCTATGT
CCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGGCTAATGGCTCAGTGTTGGCCCAGGA
GGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCTGTGCGAACACGGAAATGCCTCCAGT
AAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGGCTCAATTTAAATCATGTTCTAGT
AATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGTTCAAATGATCTCCAAGGGCCCT
TATACCCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCAAACCTAAGAACCAGGTGCA
TTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACATTTTGGGAGGCCGAGGCG
GGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCCCTGTCTC
TACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATCCCAGCTACTCGGG
AGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACAGGAGAATCACT
TCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTATGGTTATTT
GTAA

114/168

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTAS
PCWPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:

amino acids 1-15

114/168

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACC**ATCGC**
AAAGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAA
GAAAATATGTAAATCACTTAAGATTTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAAT
TGTCTGTGTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGA
GCACACTTTCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAG
AACTGAAATATTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAA
CGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGT
GATTCCTGAATTTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTT
CTTTGAACAGTCAGTGATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAA
AAATTCCAAAATTCTGGAGATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAAT
ATCAGTTTCTGAGTTACAAGACTTTGAGGAGGAGGGAGAAGATCTTCACTTTCCTGCCAACGA
AAAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCG
TCACGCCAGACAAGCAAGTGAGGAAGAACTTCCAATAAATGACTATACTGAAAATGGAATAGA
ATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGTGAGGCAACCGCTA
TTGCCGCCGCGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATACTGCTACCAAGGAGG
ACGAGTCATCTGTCGTGTCATCATGCCTTGTAAGTGGTGGGTGGCCCGCATGCTGGGGAGGGT
CTAATAGGAGGTTTGAGCTCAAATGCTTAACTGCTGGCAACATATAATAAATGCATGCTATT
CAATGAATTTCTGCCTATGAGGCATCTGGCCCTGGTAGCCAGCTCTCCAGAATTACTTGTAG
GTAATTCCTCTCTTCATGTTCTAATAAACTTCTACATTATCACCAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALT LIVLFWGSKHFWPEVPPKKAYD
MEHTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQI
KVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAEKPIENRDFLKNSKILEICDNVTMYWINPT
LISVSELQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENG
IEFDPMLDERGYCCIIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGC
AGGCGGCAGGGCGGGCGGCCAGGATC**ATG**TCCACCACCACATGCCAAGTGGTGGCGTTCTCTCC
TGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGG
ACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGA
GGCAGAGTTCAGGCTTCACCGAATGCAGGCCCTATTTACCATCCTGGGACTTCCAGCCATGC
TGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCCTGGTAT
CCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGA
CACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGT
TTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTG
GGATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGCGGCTCTGTTCTGGGGCTGGGTGC
CTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAG
AAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTGCCTACAAGCCTG
GAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAACAAGAAGATATACGATGGAG
GTGCCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGACTATGTG**TAA**TGCTCTA
AGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCACCCAAAAACAAGGAGATCCCA
TCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTTTCATCTTTG
GAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACAGCTGAG
TTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAA
CTTTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAG
ACTCCCCCTCTTCCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAG
AAAACCTTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTC
CCCACCCCCAACTTGGCTAGTAATAAACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATT
TGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAAGTTACCAAACCA
AAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTGCTGTTGACATCTTCTTATTACAGC
AACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAGTCCTCTTTCTGTGCGGGTCAGAAA
TTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAATTTAAGTCCTAAATATAGTTAAATAA
ATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAG
AAGGAAATGAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAA
GTACAAATTCCATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGG
ATCACTTGAGCCCAGAAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAA
ATACAGAGAGAAAAAATCAGCCAGTCATGGTGGCATAACCTGTAGTCCCAGCATTCCGGGAG
GCTGAGGTGGGAGGATCACTTGAGCCCAGGAGGTTGGGGCTGCAGTGAGCCATGATCACACC
ACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAATAAAAAATAAATAATGGA
ACACAGCAAGTCCTAGGAAGTAGGTTAAACTAATTCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGFTEC
RPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFI
VSGLCIAIGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGG
VMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQ
SYPSKHDYV

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGG
AGTCCAGCTGGCTAAAACATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGG
CTGTTTCTTGGTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGA
GTGTCGGCCTTCATTGAAAACAACATCGTGGTTTTTGAAAACCTTCTGGGAAGGACTGTGGATG
AATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTT
TCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGGCT
TTCATGATGGCCATCCTTGGCATGAAATGCACCAGGTGCACGGGGGACAATGAGAAGGTGAAG
GCTCACATTCTGCTGACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGCTCATCCCT
GTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAATGTTGCCCAA
AAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCACTGGTGCTGATTGTTGGA
GGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTACAGATACTCGATA
CCTTCCCATCGCACAAACCCAAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCGTCTACTCC
AGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAATGACA
AAAATCTATATTACTTTCTCAAATGGACCCCAAAGAACTTTGATTTACTGTTCTTAAGTGC
CTAATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACAGCAGAA
TGAGATATTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGG
TTCAAGCATCTACTCTTTTTATCATTTACTTCAAATGACATTGCTAAAGACTGCATTATTTT
ACTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAGTGTAACATTTATATCTCA
CATAGAGACATGCTTATATGGTTTTATTTAAAATGAAATGCCAGTCCATTACACTGAATAAAT
AGAACTCAACTATTGCTTTTCAGGGAAATCATGGATAGGGTTGAAGAAGGTTACTATTAATTG
TTTAAAAACAGCTTAGGGATTAATGTCCTCCATTTATAATGAAGATTAAAATGAAGGCTTTAA
TCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATATGCTGTTTTTTAGCCTAGGAGTTAGAA
ATCCTAACTTCTTTATCCTCTTCTCCCAGAGGCTTTTTTTTTCTTGTGTATTAAATTAACATT
TTTAAAACGCAGATATTTTGTCAAGGGGCTTTCATTCAAACCTGCTTTTCCAGGGCTATACTC
AGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTGAAAATATTTTT
GTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGATATATTT
TAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAGCAGAAAAATAT
GTCTTGGTTTTCAATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAAGTTC
ACCTGCTCCTATGTGGGTACCTGAGTCAAATTTGTCATTTTTGTTCTGTGAAAAATAAATTTT
CTTCTGTACCATTTCTGTTTAGTTTTACTAAAATCTGTAAATACTGTATTTTTCTGTTTATT
CCAAATTTGATGAACTGACAATCCAATTTGAAAGTTTGTGTCGACGTCTGTCTAGCTTAAAT
GAATGTGTTCTATTTGCTTTATACATTTATATTAATAAATTGTACATTTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQ
CKIYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIF
IITGMVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVVLIVGGALFCCVFCCN
EKSSSYRYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGG
AGCCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCGGC
AGCCGGGAGCC**ATG**CGACCCCAAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCGGCCTCCTGC
TGCTCCTGCTGCTGCAGCTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAA
AGGCGCAGCTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAG
CAGGAGTGCCTGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCC
CAGGTCGGGATGGATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCT
GGACACCCAACTACAAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAA
TTGCGGAGTGTACATTTACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCT
CACTTCGGCTAAAATGCAGAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTG
AATGTTTCAGGACCTCTTCCCATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGA
ATTCAACAATTAATATTCATCGCACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTG
GATTAGTGGATGTTGCTATCTGGGTTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTA
CTGGATGGAATTTCAGTTTCTCGCATCATTATTGAAGAACTACCAAAA**TAA**ATGCTTTAATTTT
CATTTGCTACCTCTTTTTTTTATTATGCCTTGGAATGGTTCACTTAAATGACATTTTAAATAAG
TTTATGTATACATCTGAATGAAAAGCAAAGCTAAATATGTTTACAGACCAAAGTGTGATTTCA
CACTGTTTTTAAATCTAGCATTATTCATTTTGCTTCAATCAAAAGTGGTTTCAATATTTTTTT
TAGTTGGTTAGAATACTTTCTTCATAGTCACATTCTCTCAACCTATAATTTGGAATATTGTTG
TGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTAAAAAATATAAAAGCTACCAATCTTTG
TACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGTAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPA
GVP
GRDGSPGANVIPGTPGIPGRDGFKEGEGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAEC
TFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSP
PEMNSTI
NIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

122/168

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTG
AACTGGGTGCTCATCACGGGAAGTGTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCC
CCAAATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTT
TTTTTAACCGCCCCCTCCCCACCCCCCAAAAAAAGTGTAAAGATGCAAAAACGTAATATCCAT
GAAGATCCTATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTT
GTTCTTGGAGTGTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAG
GGGTCCAATTTTTCTCCTGGGTGTGAGCGAGCCCTGACTACTACAGTGCAGCTGACAGGGG
CTGTGATGCAACTGGCCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAA
AGG**ATG**GGTTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCC
ACTGTCTTACTGACAATGCTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAA
GGCAAAATGGTATATTGTGAATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGT
TGCTTAGGTTTGTCCCTTCGCTATAACAGCCTTCAAAAAGTAAAGTATAATCAATTTAAAGGG
CTCAACCAGCTCACCTGGCTATACCTTGACCATAACCATATCAGCAATATTGACGAAAATGCT
TTTAATGGAATACGCAGACTCAAAGAGCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTT
AACAAATACCTTCAGACCTGTGACAAATTTACGGAAGTGGATCTGTCTTATAATCAGCTGCAT
TCTCTGGGATCTGAACAGTTTCGGGGCTTGCGGAAGCTGCTGAGTTTACATTTACGGTCTAAC
TCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCGCAACCTGGAACCTTTTGGACCTG
GGATATAACCGGATCCGAAGTTTAGCCAGGAATGTCTTGGCTGGCATGATCAGACTCAAAGAA
CTTCACCTGGAGCACAATCAATTTTCCAAGCTCAACCTGGCCCTTTTCCAAGGTTGGTCAGC
CTTCAGAACCTTTACTTGCACTGGAATAAAATCAGTGTATAGGACAGACCATGTCTTGGACC
TGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGATCGAAGCTTTCAGTGGACCCAGT
GTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGATTCCAACAAGCTCACATTTATT
GGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAGTCTTGCTGGGAATATATGG
GAATGCAGCAGAAATATTTGCTCCCTTGTAAGCTGGCTGAAAAGTTTTAAAGGTCTAAGGGAG
AATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTGATCGATGCAGTGAAG
AACTACAGCATCTGTGGCAAAGTACTACAGAGAGGTTTGATCTGGCCAGGGCTCTCCCAAAG
CCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTTTGCCCCCGACG
GTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCTTTCCATAAA
ATCATCGCGGGCAGCGTGGCGCTTTTCCCTGTCCGTGCTCGTCATCCTGCTGGTTATCTACGTG
TCATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAAGGCAC
AGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTTTATGTAGAT
TATAAACCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTAT
AACAAATCGGGCTCCAGGGAGTGTGAGGTAT**TGA**ACCATTGTGATAAAAAGAGCTCTTAAAGC
TGGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCTC
CCCTTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCCCTTGTCGGTTTTAGTGCATTATA
ATACTGGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGT
GAAGCTTGAAGTCCGGTTTAATATAATACCTATTGTATAAGACCCCTTACTGATTCCATTAAT
GTCGCATTTGTTTTAAGATAAAACTTCTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGC
LGLSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLN
NTERPVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELLDLG
YNRIRSLARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQONLYLQWNKISVIGQTMSTW
SSLQRLDLSGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIGQEILDSWISLNDISLAGNIWE
CSRNICSLVNWLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKP
TFKPKLPRPKHESKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVS
WKRYPASMKQLQQRSLMRRHRKKKRQSLKQMTPTSTQEFYVDYKPTNTETSEMLLNGTGPTYN
KSGSRECEV

Important features of the protein:**Signal peptide:**

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGA**ATG**TCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCGC
TGACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCG
AGCTAGCAACCTTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTC
TTGCTCGGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCC
TAGGGATCATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACA
GACACGTAGTGTATTCTGGAGGTGCAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTG
GCAAAAGTGAAGATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTG
TTATTGGCCAGTTTTTAGCCAATCCAACTGACCTAGTGAAGGTTGAGATGCAAATGGAAGGAA
AAAGGAACTGGAAGGAAAACCATTCGATTTTCGTGGTGTACATCATGCAATTTGCAAAAATCT
TAGCTGAAGGAGGAATACGAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCAC
TGGTGAATATGGGAGATTTAACCCTTATGATACAGTGAAACACTACTTGGTATTGAATACAC
CACTTGAGGACAATATCATGACTCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTA
TTCTGGGAACACCAGCCGATGTCATCAAAGCAGAATAATGAATCAACCACGAGATAACAAG
GAAGGGGACTTTTGTATAAATCATCGACTGACTGCTTGATTCAGGCTGTTCAAGGTGAAGGAT
TCATGAGTCTATATAAAGGCTTTTTACCATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGT
TCTGGCTTACTTATGAAAAAATCAGAGAGATGAGTGGAGTCAGTCCATTT**TAA**

FIGURE 126

MSVPEEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGAR
ESAPYRGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEHYP
LWKSVIGGMMAGVIGQFLANPTDLVKVQMOMEKGRKLEGKPLRFRGVHHAFAKILAEGGIRGL
WAGWVPNIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVI
KSRIMNQPRDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIR
EMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

126/168

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCG
ACCCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGGCGTGGGCCC**ATG**GCCAGGCCCGGC
ATGGAGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCCTCGGGGGGCATCGGCGCGGCC
GTGGCCCCGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAAC
ATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCTACAGA
TGTGACCTATCAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGC
GGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCCGGCCTGACACCCTGCTCTCAGGCAGC
ACCA GTGGTTGGAAGGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAA
GCCTACCAGTCCATGAAGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATG
TCTGGCCACCGAGTGTTACCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTC
ACTGCGCTGACAGAGGGACTGAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCACG
TGCATCTCTCCAGGTGTGGTGGAGACACAATTCGCCTTCAAACCTCCACGACAAGGACCCTGAG
AAGGCAGCTGCCACCTATGAGCAAATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTT
ATCTACGTCCTCAGCACCCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACGGAG
CAGGTGACC**TAG**TGACTGTGGGAGCTCCTCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTG
CCTCTGGATTTTAGGTGTTGATTTCTGGATCACGGGATACCACTTCCTGTCCACACCCCGACC
AGGGGCTAGAAAATTTGTTTGAGATTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATG
TGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCCTAATTGTTTTACTTGTTAACTTGTTCTTG
TGCCCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTT
GTGGCCAAAATCCCCATCTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG
GAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCCTCTGCCTGCCCC
ACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCCAGCCAGTCTTGCTTCTTGT
CCCCCTCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCAGGGCCTGGC
CCAGTGGATTTTCATGGTGATCATTA AAAAAGAAAAATCGCAACCAAAAAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGA A VARALVQQGLKVVG CARTVGNIEELAAECKSAGY PGT
LIPYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFN VNVLALS
ICTREAYQSMKERNVDDGHIININSMGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQT
HIRATCISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQ
MRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTAC**ATG**GGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACC
ATCATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTTCAGGTGCAGAGTCTCAGTTGCCCGG
GAGCACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATTCCAGTTCTGGTT
TCATGCCAGCCTGTAAAAGGCCATGGAACCTTTGGGTGAATCACCGATGCCATTTAAGAGGGTT
TTCTGCCAGGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAG
CCACCTGTGGCCGTTGAGTGCTTGAA**TGA**AGGAAGTGAAGAAATTAATTTCTCATGTATTTTT
CTCATTTATTTATTAATTTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGG
ATACATGTATACAATATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACAT
TTATTTTTTTATTCTTTTTTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCC
ATCTCAGCTTACTGCAACCTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAA
GTAGCTGGGACTACAGGCATGCACCACAATGCCCAACTAATTTTTGTATTTTTAGTAGAGACG
GGGTTTTGCCATGTTGCCCAGGCTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCG
GCCTCCCAAAGTGTTATGATTACAGGCGTGAGCCACCGTGCTGGCCTAAACATTTATCTTTT
CTTTGTGTTGGGAACCTTTGAAATTATACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTG
CTATGGAACACTGGGACTTCTTCCCTCTATCTAACTGTATATTTGTACCAGTTAACCAACCGT
ACTTCATCCCCACTCCTCTCTATCCTTCCCAACCTCTGATCACCTCATTCTACTCTCTACCTC
CATGAGATCCACTTTTTTTAGCTCCCACATGTGAGTAAGAAAATGCAATATTTGTCTTTCTGTG
CCTGGCTTATTTCACTTAACATAATGACTTCCTGTTCCATCCATGTTGCTGCAAATGACAGGA
TTTCGTTCTTAATTTCAATTAAAATAACCACACATGGCAAAAA

130/168

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRI PVLVSCQ
PVKGHGTLGESPMPPFKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

130/168
MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRI PVLVSCQ
PVKGHGTLGESPMPPFKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAG
 CCTGGAGCATCTTCCTCATCGGGACTAAAATTGGGCTGTTTCCTTCAAGTAGCACCTCTATCAG
 TTATGGCTAAATCCTGTCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTTACTGTAATGATC
 GCTTTCTGACATCCATTCCAACAGGAATACCAGAGGATGCTACAACTCTCTACCTTCAGAACA
 ACCAAATAAATAATGCTGGGATTCCTTCAGATTTGAAAACTTGCTGAAAGTAGAAAGAATAT
 ACCTATACCACAACAGTTTAGATGAATTTCTTACCAACCTCCCAAAGTATGTAAAAGAGTTAC
 ATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAAAAATTCCTTATCTGG
 AAGAATTACATTTAGATGACAACTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCATTCCGAG
 ACAGCAACTATCTCCGACTGCTTTTCTGTCCCGTAATCACCTTAGCACAAATTCCTGGGGTT
 TGCCCAGGACTATAGAAGAACTACGCTTGATGATAATCGCATATCCACTATTTTCATCACCAT
 CTCTTCAAGGTCTCACTAGTCTAAAACGCCTGGTTCTAGATGGAAACCTGTTGAACAATCATG
 GTTTAGGTGACAAAGTTTTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATT
 CCCTGACTGCTGCACCAGTAAACCTTCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATA
 ACCACATCAATCGGGTGCCCCCAAATGCTTTTTCTTATCTAAGGCAGCTCTATCGACTGGATA
 TGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTTGATGATTTGGACAATATAACAC
 AACTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATGGGTACGTGACTGGT
 TACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAAAGGTTC
 GTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAA
 GCACCATTGAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCCAAGGACAGTGGCCAG
 CTCCAGTGACCAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAG
 GGAGTCCCTCAAGAAAAACAATTACAATTACTGTGAAGTCTGTCACCTCTGATACCATTGATA
 TCTCTTGGAACTTGCTCTACCTATGACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATA
 GCCCGGCATTTGGATCTATAACAGAAACAATTGTAACAGGGGAACGCAGTGAGTACTTGGTCA
 CAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTCCCATGGAAACCAGCAACCTCT
 ACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTTCGAATGTACAACC
 CTACAACCACCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACCTTTGG
 CTGCCATCATTTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTTGGT
 ATGTTTCATAGGAATGGATCGCTCTTCTCAAGGAAGTGTGCATATAGCAAAGGGAGGAGAAGAA
 AGGATGACTATGCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAACTT
 CTTTTCAGATGTTACCAATAAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAATACACACCA
 TATTTCTCCTAATGGAATGAATCTGTACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAA
 GCTACAGAGACAGTGGTATTCCAGACTCAGATCACTCACACTCATGATGCTGAAGGACTCACA
 GCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATGGT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCND RFLTSIPTGIPEDATTL
 YLQNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNI RTITYDSL SK
 IPYLEELHLDDNSVSAVSIIEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRIST
 ISSPSLQGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKL
 YLQDNHINRVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDLDNITQLILRNNPWYCGCKMKW
 VRDWLQSLPVKVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQ
 GQWPAPVTKQPDIKNP KLT KDQQT TGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWL
 KLGHS PAFGSITETIVTGERSEYLVTALEPDSPYKVCMPMETSNLYLFDETPVC IETETAPL
 RMYNP TTTLNREQEKEPYKNPNLPLAAIIGGAVALVTIAL LALVCWYVHRNGSLFSRNCAYSK
 GRRRKDDYAEAGTKKDNSILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNHSES
 SSNRSYRDSGIPDSDHSHS

Important features of the protein:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
 640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCAGGCCACCCCAGGCTTCTTGG
CAGCCCTGCCGGGCCACTTGTCTTC**ATG**TCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAG
GGCGTGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGA
GCAGAGGCCCGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACC
GCTGTGGTCCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCT
GATCCTGTCTGCGCTGGAGAGAGCCACCGTCTTCCTAGAACAGAGGCTGCCTGAAATCAACCT
GGATGGCATGGTGGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGC
CCAGGAGCCCCCTGCTGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGC
TGCCATCCAGAGATCCCTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCA
GCTGACCCTCCAGCCCGGGTTTTTGGAAGCTCCCACATGCCTGGATCCACACTGATGCCTCCTT
GGTGTAACCCACGTTTCGGGCCCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGT
GCAGCTGCTGGGAACCGGGACGGACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAG
CCTCATGACCAAGCCCGGCTGCTCAGGCTACTGCCTGTCCCACCAACTGCTCTTCTTCCTCTG
GGCCAGAATGAGGGGATGCACACAGGGACCACTCCAACAGAGCCAGGACTATATCAACCTCTT
CTGCGCCAACATGATGGACTTGAACCGCAGAGCTGAGGCCATCGGATACGCCTACCCTACCCG
GGACATCTTCATGGAAAACATCATGTTCTGTGGAATGGGCGGCTTCTCCGACTTCTACAAGCT
CCGGTGGCTGGAGGCCATTCTCAGCTGGCAGAAACAGCAGGAAGGATGCTTCGGGGAGCCTGA
TGCTGAAGATGAAGAATTATCTAAAGCTATTCAATATCAGCAGCATTTTTTCGAGGAGAGTGAA
GAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGTTGCTCAGGCTGGAGTACAGTGGCGCAA
TCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGCAATTCTCTTGCCTCATCCTCCCGAG
TAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAATTTTTATATTTTTTTTAGTAGAGAC
AGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGATCTCAAGAGATCCGCCCACCTC
AGGCTCCCAAAGTGTTGGGATT**TAG**GTGTGAGCCACCGTGTCTGGCTGAAAAGCACTTTCAAA
GAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTCATGGGGGCTCTCTCCCC
TAGATGGCTGCTCCTCCCACAACACAGCCACAGCAGTGGCAGCCCTGGGTGGCTTCCTATACA
TCCTGGCAGAATACCCCCCAGCAAACAGAGAGCCACACCCATCCACACCGCCACCACCAAGCA
GCCGCTGAGACGGACGGTTCATGCCAGCTGCCTGGAGGAGGAACAGACCCCTTTAGTCCTCA
TCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGGATAAGCAAA
GCCACCCCGACACCCAATCTTGGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGGGGCCGGG
AGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLLARAQGAERVTSSEQRPAMASLGLLLLLLLLTALPPLWSSSLPG
 LDTAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPL
 SLRVGMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQ
 DSFSEERSDVCLVQLLGTGTDSSEPCGLSDLCSLMTKPGCSGYCLSHQLLFFLWARMRGCTQ
 GPLQQSQDYINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILS
 WQKQQEGCFGEPPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLP
 PGFKQFSLILPSSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTC**ATG**GCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGC
TGCTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCCGGGGCTGCTGCTGAGGGATCGG
GAGGGAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAGTTGTTCCAGGGG
TGAAGCCTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTT
TCCTTAAGACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAG
TTGTATCTCCAGCTTACAGATTTGATCCCGTTCGAGTGATATCACTTCGAAAGGAAAAATGA
GAGCAAGATATGTGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAA
TGAAATCTTCAGGTCCACCTTCTTACTTTATTAAAAGGGAATCGTGGGGCTGGACAGACTTTC
TAATGAACCCAATGGTTATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAG
TGGTCAACACAAGTGATCCTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATT
CCAACCATGAGTTGCCTGATGTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTG
GCAAATCTAGCAGCGGCAGCAGTAAACAGGCAAAAGTGGGGCTGGCAAAAGGAGG**TAG**TCAG
GCCGTCCAGAGCTGGCATTTCACAAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAAA
CCGTGTGAAGCAACTACTATAAACTTGAGTCATCCCGACGTTGATCTCTTACAACGTGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGT
ATGAGGTCAATATTGATGTCAGTGAATTAATTACAGTGTCCTATAGAAAATGCCATTAATAAA
TTATATGAACTACTATACATTATGTATATTAATTAAAACATCTTAATCCAGAAATCAAAAAA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGSGGSGVGIGDRFKIEGRAVVPGVKPQDWIS
AARVLVDGEEHVGFLKTDGSFVVHDIPSGSYVVEVVSPAYRFDVVRVDITSKGKMRARYVNYI
KTSEVVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMMLVLLIFVLLPKVVNTSDPD
MRREMEQSMNMLNSNHELDPVSEFMTRLFSSKSSGKSSSGSSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGG
GACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAA
ACAAGTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTT
CCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACG
ACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAG
GCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTACGGCCTAAGATGAAAGCCTCTAGT
CTTGCCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAG
ACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCT
GAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACT
GAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTC
TATCTGGACAGGGTATTTAAAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGC
AGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACA
TGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTG
GAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAG
GAGACAGAATAGGAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCT
TCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACTGTACTAGTCTTGTGCT
GGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGATTGTCTTTATGCATCCCC
AATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAAATATCTTTCTGCTATTGGATATATT
TATTAGTTAATATATTTATTTATTTTTTGTCTATTTAATGTATTTATTTTTTTTACTTGGACATG
AACTTTAAAAAAATTACAGATTATATTTATAACCTGACTAGAGCAGGTGATGTATTTTTTAT
ACAGTAAAAAAAACCTTGTAATTCTAGAAGAGTGGCTAGGGGGGTATTTCATTTGTAT
TCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGATATTTGAAATTGAACCAATGAC
TACTTAGGATGGGTTGTGGAATAAGTTTTGATGTGGAATTGCACATCTACCTTACAATTACTG
ACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAGCCAGGAATCCTACACGG
CCAGCATGTATTTCTACAAATAAAGTTTTCTTGCATACCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNGATTTCATNSHSDSELRPE
IFSSREAWQFFLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQE
IRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHY
TLRKISSLANSLFTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAADVVKALGELDI
LLQWMEETE

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCCGCATCCAG
CCTAGCGTGTCCACG**ATG**CGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAG
GTGCTAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCCTCTGCCAGA
GCGGAACACGGAGCGGAGCCCCAGCGCCCGAACCCTCGGCTGGAGCCAGTTCTAACTGGACC
ACGCTGCCACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGAT
TTTGTGTTTGGGTCAAAGGGTGTGAAATTTATGCCCTACACAACTTACCTTGTGGAAAAAGGA
GCATCTCACAGTTTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCA
TTGATGACGGGGAGCCTTCCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTCCTGCACTG
CTGGAAGACAGTGTGATAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGAT
GAAACCTGGGTAAATTATTCCCAAAGCATTTTGTGGAATATGATGGAACAACCTCATTTTTC
GTGTCAGATTACACAGAGGTGGATAATAATGTCACGAGGCATTTGGATAAAGTATTA AAAAGA
GGAGATTGGGACATATTAATCCTCCACTACCTGGGGCTGGACCACATTGGCCACATTTTCAGGG
CCCAACAGCCCCCTGATTGGGCAGAAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCAC
ACCTCACTGCAGTCGAAGGAGAGAGAGACGCCTTTACCCAATTTGCTGGTCTTTTGTGGTGAC
CATGGCATGTCTGAAACAGGAAGTCACGGGGCCTCCTCCACCGAGGAGGTGAATACACCTCTG
ATTTTAATCAGTTCTGCGTTTGAAAGGAAACCCGGTGATATCCGACATCCAAAGCACGTCCAA
TAGACGGATGTGGCTGCGACACTGGCGATAGCACTTGGCTTACCGATTCCAAAAGACAGTGTA
GGGAGCCTCCTATTCCAGTTGTGGAAGGAAGACCAATGAGAGAGCAGTTGAGATTTTTACAT
TTGAATACAGTGCAGCTTAGTAACTGTTGCAAGAGAATGTGCCGTATATGAAAAAGATCCT
GGGTTTGAGCAGTTTAAAATGTGAGAAAGATTGCATGGGAACTGGATCAGACTGTACTTGGAG
GAAAAGCATTCAGAAGTCCTATTCAACCTGGGCTCCAAGGTTCTCAGGCAGTACCTGGATGCT
CTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCCAGTTCTCACCTGCTCCTGCTCA
GCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCACTGTCTCCTGGGTTTT
CTCTGCTCTTTTATTTGGTGATCCTGGTTCTTTTCGGCCGTTACAGTCATTGTGTGCACCTCAG
CTGAAAGTTCGTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCTGCCTTTCGTTTACCA
GACTCTGGTTGAACACCTGGTGTGTGCCAAGTGCTGGCAGTGCCCTGGACAGGGGGCCTCAGG
GAAGGACGTGGAGCAGCCTTATCCCAGGCCTCTGGGTGTCCCGACACAGGTGTTACATCTGT
GCTGTCAGGTCAGATGCCTCAGTTCTTGGAAGCTAGGTTCTCTGCGACTGTTACCAAGGTGAT
TGTAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAATCGGACA
GCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGACACTC
AGGAGGGTCAAAGGAGACTTGGTTCGACCACTCATCCTGCCACCCCAAGATGCATCCTGCC
TCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTCTGTTGGAATTCTTAGTCCTTGGCCT
CGGACACCTTCATTTCGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCA
CACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCC
CCCACCCCAACCCTGCACAGCCCTCATCCCTCTTGCTTGAGCCGTGAGAGGCCCTGTGCTG
AGTGTCTGACCGAGACACTCACAGCTTTGTGTCATCAGGGCACAGGCTTCCTCGGAGCCAGGATG
ATCTGTGCCACGCTTGACCTCGGGCCCATCTGGGCTCATGCTCTCTCCTGCTATTGAATT
AGTACCTAGCTGCACACAGTATGTAGTTACCAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEHGAEPPEPSAGASSNWTTLPPP
LFSKVIVVLIDALRDDFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGS
LPGFVDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPHFVEYDGTTSFFVSDYT
EVDNNVTRHLDKVLKRGDWDILILHYLGLDHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQS
KERETPLPNLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTA
CAGAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTG**ATG**TTACTGCTGCTGTTGGAGTAC
AACTTCCCTATAGAAAACAACCTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAAC
TTAAACCCGAAGAAATTCAGCATTGATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGG
AATCTCATAGCAGTTCCAGATAAAAACTACATACGCCCCAGAGATCTTCTTTGCATTAGCCTCA
TCCTTGAGCTCAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAG
TTTTGTCTCTACTGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAG
AAACTGATGAAGCTGGCTGCCCCAAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCT
CAGGTGGGCTCCTGGAACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCC
TGCAATTGTAATGAGCCTGTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTT
TCATTTCAACCAGTTTGCAAAGCTGAAATGAGCCCCAGTGAGGTCAGCGAT**TAG**GAAACTGCC
CCATTGAACGCCTTCCTCGCTAATTTGAACTAATTGTATAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLDSGNLIAVPDKNYIRP
EIFFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESAR
RPFIFYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPS
EVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622

AAAAAAAAAAAAAAAA

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSK
DYYAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTI
STRPWMTQFSLLNKTCLGEGH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCG**ATG**AGGA
AGCACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGG
TCCAGACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCCAGCACTGCCC
AGATCACTGAGGCCCAGGTGGCTGAGAACCGCCCGGGAGCCTTCATCAAGCAAGGCCGCAAGC
TCGACATTGACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCG
ATGGCATCCACTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTACCCGGCT
GCATCAATGCCACCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACC
AGCAGGTGCTCTGGCGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTTGGTTGG
AGAGGGGCGCAGGACTTCGGGTCACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGA
TCTGGCTCATGGTGAAA**TAA**GCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAA
TCCTGGCAAGTGACCCAGCTCTTCTCCCCCAAACCCACGCGTGTTCTGAAGGTGCCCAGGAGC
GGCGATGCACTCGCACTGCAAATGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCT
GATAGATGGGGGACTGTGGCTTCTCCGTCACTCCATTCTCAGCCCCCTAGCAGAGCGTCTGGCA
CACTAGATTAGTAGTAAATGCTTGATGAGAAGAACACATCAGGCACTGCGCCACCTGCTTCAC
AGTACTTCCCAACAACCTCTTAGAGGTAGGTGTATTCCCGTTTTACAGATAAGGAACTGAGGC
CCAGAGAGCTGAAGTACTGCACCCAGCATCACCAGCTAGAAAGTGGCAGAGCCAGGATTCAAC
CCTGGCTTGTCTAACCCCAGGTTTTCTGCTCTGTCCAATTCCAGAGCTGTCTGGTGATCACTT
TATGTCTCACAGGGACCCACATCCAAACATGTATCTCTAATGAAATTGTGAAAGCTCCATGTT
TAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQG
RKLDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNK
LHQQVLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWLMLVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCAT
· TTCCTGATGATTTATAGACTCAAAGAAAACTATGTTCAGAAGCTCTCTTCTCTTCTGGCCTC
CTCTCTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCA
AGCTGGAAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCAGTCTTAGTTTTTATTTTT
TGAAATTTCAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCAT
GATGCTGAGGTTTGGGGT

147/168

FIGURE 148

MFRSSLLFWPPLCLLSLFLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGT
CEGLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

148/168

FIGURE 149

GTCTCCGCGTCACAGGAACTTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGAC
TTGACTCCCGCGCGCCCCAACCCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGC
CGCCAGTCCCGGCCCTCTCCCGCCCCACACCCACCCTCCTGGCTCTTCCTGTTTTTACTCC
TCCTTTTCATTACATAACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGGCTGTGACCCAAGCC
GAGCGTGGAAGAATGGGGTTCCTCGGGACCGGCACTTGGATTCTGGTGTTAGTGCTCCCGATT
CAAGCTTTCCCCAAACCTGGAGGAAGCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCA
GAAAGACCTTTGAATGAACAGATTGCTGAAGCAGAAGAAGACAAGATTAAAAAACATATCCT
CCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTTTGTGATAACTTGAACTGCTAAAGGCA
ATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTATAAGAAGCTCCCCACTTGATAAT
AAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACCTGATCGATGATTATGACTCT
ACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCATCAACTAGACGGG
ACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAGAAAATGAC
AGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCACAGAAAGCCAA
GCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAAATTAATCTCAAAGGAAGCCAAC
AATTATGAGGAGGATCCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCA
GAGAAAGTGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACA
GTATCTAACACATTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGAC
AACTTTGAGGAACTCCAATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAA
AAAGAAGCAAAAGAGAAAGAAACACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAG
ATGATGGTGAAATATGGAACAATATCTCCAGAAGAAGGTGTTTCCTACCTTGAAAACCTTGAT
GAAATGATTGCTCTTCAGACCAAAAACAAGCTAGAAAAAATGCTACTGACAATATAAGCAAG
CTTTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAACAGACAGTACCAAGGAAGAAGCAGCT
AAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAAGATGATAACTCCAACCCAGGA
GGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGCCATCAGAAAAAATATT
GAATGGTTGAAGAAACATGACAAAAAGGGAATAAAGAAGATTATGACCTTTCAAAGATGAGA
GACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAAGAAGCC
GAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAATGGCAAAGATCCAGGAGTCTTTCAAC
TGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATTCTGTGATTAAATTTTTTGACCC
AAGGGTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGC
TTTCTTCCCGTAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAA
AAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENK
PGQSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSG
LDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTL
EDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVT PMAAIQDGLAKGENDETVSNT
LTLTNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVK
YGTISPEEGVSYLENLDEMIALQTKNKLEKNATDNISKLF PAPSEKSHEETDSTKEEAAKMEK
EYGSLKDSTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFIN
KQADAYVEKGILDKEEA EAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTC
 AAG**ATG**GTTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTAT
 CTGCATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAG
 ATCAGCGTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCCGTCATCCTGGGTGTCCAG
 GGTGGAAGCCAGTGCCTGTGCATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTG
 AACATCATGGAGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTCACCTTCTACCGGCGGGAC
 ATGGGGCTCACCTCCAGCTTCGAGTCGGCTGCCTACCGGGCTGGTTCTGTGCACGGTGCCT
 GAAGCCGATCAGCCTGTGAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATC
 ACAGACTTCTACTTCCAGCAGTGTGACT**TAG**GGCAACGTGCCCCCAGAACTCCCTGGGCAGAG
 CCAGCTCGGGTGAGGGGTGAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAG
 GACCCCACTGCTGACTTAGTGGGCACCTGACCACTTTGTCTTCTGGTTCCAGTTTGATAA
 ATTCTGAGATTTGGAGCTCAGTCCACGGTCTCCCCACTGGATGGTGTCTGCTGTGGAAC
 CTTGTAAAAACCATGTGGGGTAAACTGGGAATAACATGAAAAGATTTCTGTGGGGGTGGGGTG
 GGGGAGTGGTGGGAATCATTCTGTCTAATGGTAACTGACAAGTGTACCCTGAGCCCCGAG
 GCCAACCCATCCCCAGTTGAGCCTTATAGGGTCAGTAGCTCTCCACATGAAGTCTGTCACTC
 ACCACTGTGCAGGAGAGGGAGGTGGTCATAGAGTCAGGGATCTATGGCCCTTGCCCGAGCCCC
 ACCCCCTTCCCTTTAATCCTGCCACTGTCTATGTACCTTTCCCTATCTCTTCCCTCATCATC
 TTGTTGTGGGCATGAGGAGGTGGTGATGTCAGAAGAAATGGCTCGAGCTCAGAAGATAAAAGA
 TAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAAGATAACAATCAAATCCCAGATGCTGGT
 CTCTATTCCCATGAAAAAGTGCTCATGACATATTGAGAAGACCTACTTACAAAGTGGCATATA
 TTGCAATTTATTTTAAATTAAGATAACCTATTTATATATTTCTTTATAGAAAAAGTCTGGAA
 GAGTTTACTTCAATTGTAGCAATGTCAGGGTGGTGGCAGTATAGGTGATTTTTCTTTTAATTC
 TGTTAATTTATCTGTATTTCTTAATTTTTCTACAATGAAGATGAATTCCTTGTATAAAAAATAA
 GAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGATTGTCCTCAGCCTCCAC
 TTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGGTTGTAGTAGTGATCA
 GGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGTGTGGCTGGAATCT
 CTGGGTAAGGAACCTTAAAGAACAAAAATCATCTGGTAATTCTTTCCCTAGAAGGATCACAGCCC
 CTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGAATTGTGTCC
 CCTCAAATTCACATCCTTCTTGGAAATCTCAGTCTGTGAGTTTATTTGGAGATAAGGTCTCTG
 CAGATGTAGTTAGTTAAGACAAGGTCATGCTGGATGAAGGTAGACCTAAATTCAATATGACTG
 GTTTCCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTAAAG
 ATGAAGGCAGAGATCGGAGTTTTGCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACCA
 TCAGAAGCTTGGAAGAGGCAAAGAAGAATTCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCT
 GCTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGAATAAATTTGGCTGTTT
 TAAGCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATC
 CCTGTCTCCTCGTGTTTACATTCTGTGTGTGTCCCCCTCCACAATGTACCAAAGTTGTCTTTG
 TGACCAATAGAAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACAC
 TGCAGCTTCTACTTGAGCCCTCTCTCTCTGCCACCCACCGCCCCCAATCTATCTTGGCTCACT
 CGCTCTGGGGGAAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAA
 AATGAAGTCTCCTGCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATC
 GATGTTTGTGTTTAAAGTTGCTCAGTTTTGGTCTAACTTGTTATGCAGCAATAGATAAATAA
 TATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESISVVPNRWLDASLSPVILGVQG
GSQCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPE
ADQPVRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCA**ATGGCC**
GCCCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTC
TTGGCCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCAGTGCAGGCTTGACAAG
TCCAACCTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTG
GCTGATAACAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGT
GAGCGCTGCTATCTGATGAAGCAGGTGCTGAACTTCACCCTTGAAGAAGTGCTGTTCCCTCAA
TCTGATAGGTTCCAGCCTTATATGCAGGAGGTGGTGCCCTTCCTGGCCAGGCTCAGCAACAGG
CTAAGCACATGTCATATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAG
GACACAGTGAAAAAGCTTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAAGTGGATTTGCTG
TTTATGTCTCTGAGAAATGCCTGCATT**TGAC**CCAGAGCAAAGCTGAAAAATGAATAACTAACCC
CCTTTCCTGCTAGAAATAACAATTAGATGCCCCAAAGCGATTTTTTTTAAACCAAAGGAAGA
TGGGAAGCCAAACTCCATCATGATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGG
AAACCAATGCCACTTTTGTTTATAAGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTG
ATAACATTTTCATTGTAAGTGGTGTCTATACACAGAAAACAATTTATTTTTTAAATAATTGTC
TTTTTCCATAAAAAAGATTACTTTCCATTCCCTTAGGGGAAAAAACCCCTAAATAGCTTCATG
TTTCCATAATCAGTACTTTATATTTATAAATGTATTTATTATTATTATAAGACTGCATTTTAT
TTATATCATTTTATTAATATGGATTTATTTATAGAAACATCATTCGATATTGCTACTTGAGTG
TAAGGCTAATATTGATATTTATGACAATAATTATAGAGCTATAACATGTTTATTTGACCTCAA
TAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEA
SLADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLS
NRLSTCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCT
GTCAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTCA GTGCAGA
GGGCTGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGG
CCAAGCTGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGG
GTGCAGATGAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCCTACAG
GTGGTTGCATTCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGC
CCCAGCAAAGGGCAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCC
CTAGAGCCTGCTAGGCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTC
AACAGCAGGGCCATCTCCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAG
GACCTGTACCACGCCC GTTGCCTGTGCCCGCACTGCGTCAGCCTACAGACAGGCTCCACATG
GACCCCCGGGGCAACTCGGAGCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGC
CATGGCGAGAAGGGCACCCACAAGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTTCCTTA
GCTTGTGTGTGTGTGCGGCCCCGTGTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTT
TTGGGAAACCTGGAGCCAGGTGTACAACCACTTGCCATGAAGGGCCAGGATGCCCAGATGCTT
GGCCCCCTGTGAAGTGCTGTCTGGAGCAGCAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAA
AACCTGCACTTCTGCACATTTTGAAAAGAGCAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGT
CCTGTCATTTTCTCTCAGGAAAGGTTTTCAAAGTTCTGCCCATTTCTGGAGGCCACCACTCCT
GTCTCTTCCTCTTTTCCCATCCCCTGCTACCCTGGCCCAGCACAGGCACTTTCTAGATATTTC
CCCCTTGCTGGAGAAGAAAGAGCCCCTGGTTTTATTTGTTTGTTTACTCATCACTCAGTGAGC
ATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTTTTGACATGGATGATTCTGAGGAGGA
AGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATATCTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGTHYSHWPSCCPSKGQDTSEELLRWSTVPVPPLE
PARPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDP
RGNSELLYHNQTVFYRRPCHGEKGTGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAG
CCGACCGTTCAATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTA
ATCCCCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAAC TAGTGTTGCAACAGGGGAC
TATTCAATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAG
GCCACCAAGATTTGTGTGACGGGCAAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAAT
TACACAGAGGCCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCCTACATC
GGCTTCCCTGTAGAGCTGAACACAGTCTATTTCAATTGGGGCCATAATATTCCTAATGCAAAT
ATGAATGAAGATGGCCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATA
ATGAAATATAAAAAAAGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGT
AAGAAGAATGAGGAGACAGTAGAAGTGAACCTCACAACCACTCCCCTGGGAAACAGATACATG
GCTCTTATCCAACACAGCACTATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAA
CAAACGCGAGCTTCAGTGGTGATTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTG
ACTCCATATTTTCCTACTTGTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGC
CCACAAACAGGCGTCCCTTTCCCTCTGGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCT
CTCCTCCTGCTGTCTCTGCTGGTGGCCACATGGGTGCTGGTGGCAGGGATCTATCTAATGTGG
AGGCACGAAAGGATCAAGAAGACTTCCTTTTCTACCACCACACTACTGCCCCCATTAAGGTT
CTTGTGGTTTACCCATCTGAAATATGTTTCCATCACACAATTTGTTACTTCACTGAATTTCTT
CAAAACCATTCAGAAAGTGAGGTCATCCTTGAAAAGTGGCAGAAAAAGAAAATAGCAGAGATG
GGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTCTTTCC
AATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCAGTGAGAAC
TCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAAGCCAGATTCAT
CTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACGATTACAATGCTCTC
AGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTTCTGTGCAGAACTTCTCCAT
GTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACGATGGCTGCTGCTCCTTG

TAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLI PGDLRLDLRVEPVTTTSVATGDYS
 ILMNVSWVLRADASI RLLKATKICVTGKS NFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGF
 PVELNTVYFIGAHNIPNANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKK
 NEETVEVNFTTTPLGNRYMALIQHSTIIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTP
 YFPTCGSDCIRHKGTVVLC PQTGVPFPLDNNKSKPGGWLPLLLLSLLVATWVLVAGIYLMWRH
 ERIKKT SFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFLQNHCRSEVILEKWQKKKIAEMGP
 VQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSPSENSQDLFPLAFNLFCSDLRSQIHLH
 KYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHV KQQVSAGKRSQACHDGCCSL

Important features of the protein:**Signal peptide:**

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201
 and 283 - 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

AGCCACCAGCGCAAC**ATG**ACAGTGAAGACCCCTGCATGGCCCCAGCCATGGTCAAGTACTTGCTG
CTGTTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGA
CATACTTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGAC
ATTGGCATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACC
TCCCCCTGGAATTACACTGTCACTTGGGACCCCCAACCGGTACCCCTCGGAAGTTGTACAGGCC
CAGTGTAGGAACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTT
CCCATCCAGCAAGAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTTCCAG
TTGGAGAAGGTGCTGGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAG
TAAGAGGTGCATATCCACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIIN
ENQRVMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQE
TLVVRRKHQGCSVSFQLEKVLVTVGCTCVTPVHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCA
 GGACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGG
 ATTCAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCCTTGGGGGGGGG
 CAGCACAGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAG**ATG**CCTGTGCCCTGGTTCTT
 GCTGTCCTTGGCACTGGGCCGAAGCCCAGTGGTCCTTTCTCTGGAGAGGCTTGTGGGGCCTCA
 GGACGCTACCCACTGCTCTCCGGGCCTCTCCTGCCGCCTCTGGGACAGTGACATACTCTGCCT
 GCCTGGGGACATCGTGCTGCTCCGGGGCCCCGTGCTGGCGCCTACGCACCTGCAGACAGAGCT
 GGTGCTGAGGTGCCAGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGC
 CGTGCAATGGGCACTGGGAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGG
 GGTGGAGGAGCCTAGGAATGCCTCTCTCCAGGCCCAAGTCGTGCTCTCCTTCCAGGCCTACCC
 TACTGCCCGCTGCGTCTCTGCTGGAGGTGCAAGTGCCTGCTGCCCTTGTGCAGTTTGGTCAGTC
 TGTGGGCTCTGTGGTATATGACTGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTC
 CTATACTCAGCCCAGGTACGAGAAGGAACTCAACCACACACAGCAGCTGCCTGCCCTGCCCTG
 GCTCAACGTGTCAGCAGATGGTGACAACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCA
 CTTCCGGCCTCTCCCTGTACTGGAATCAGGTCCAGGGCCCCCCTTCCAGGCCTACCC
 CCTGACTGGACCGCAGATCATTACCTTGAACCACACAGACCTGGTTCCCTGCCTCTGTATTCA
 GGTGTGGCCTCTGGAACCTGACTCCGTTAGGACGAACATCTGCCCTTTCAGGGAGGACCCCCG
 CGCACACCAGAACCTCTGGCAAGCCGCCGACTGCGACTGCTGACCCTGCAGAGCTGGCTGCT
 GGACGCACCGTGCTCGCTGCCCGCAGAAGCGGCACTGTGCTGGCGGGCTCCGGGTGGGGACCC
 CTGCCAGCCACTGGTCCCACCGCTTTCCTGGGAGAACGTCACTGTGGACAAGGTTCTCGAGTT
 CCCATTGCTGAAAGGCCACCCTAACCTCTGTGTTTCAGGTGAACAGCTCGGAGAAGCTGCAGCT
 GCAGGAGTCTTGTGGGCTGACTCCCTGGGGCCTCTCAAAGACGATGTGCTACTGTTGGAGAC
 ACAGAGGCCCCAGGACAACAGATCCCTCTGTGCCTTGGAAACCCAGTGGCTGTACTTCACTACC
 CAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTACAAGACCTGCAGTCAGG
 CCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCCTGCCCCATGGACAAATA
 CATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTTGCCGCTGCGCTTTCCCT
 CATCCTCCTTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAAACAGGACGTCCG
 CTCGGGGGCGGCCGCCAGGGGCCGCGCGCTCTGCTCCTCTACTCAGCCGATGACTCGGGTTT
 CGAGCGCCTGGTGGGCGCCCTGGCGTCCGGCCTGTGCCAGCTGCCGCTGCGCGTGGCCGTAGA
 CCTGTGGAGCCGTCGTGAACCTGAGCGCGCAGGGGCCCGTGGCTTGGTTTACGCGCAGCGGCG
 CCAGACCCTGCAGGAGGGCGGCGTGGTGGTCTTGTCTTCTCTCCGGTGCGGTGGCGCTGTG
 CAGCGAGTGGCTACAGGATGGGGTGTCCGGGCCCGGGCGCACGGCCCGCACGACGCTTCCG
 CGCCTCGCTCAGCTGCGTGTGCCGACTTCTTGACAGGGCCGGGCGCCCGGCAGCTACGTGGG
 GGCTGCTTCGACAGGCTGCTCCACCCGGACGCGGTACCCGCCCTTTTCCGCACCGTGCCCGT
 CTTCACTGCCCCTCCCAACTGCCAGACTTCCTGGGGCCCTGCAGCAGCCTCGCGCCCCGCG
 TTCCGGGGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTCAGCCAGCCCTGGATAG
 CTACTTCCATCCCCCGGGGACTCCCGCGCCGGGACGCGGGGTGGGACCAGGGGCGGGACCTGG
 GCGGGGGACGGGACT**TAA**ATAAAGGCAGACGCTGTTTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSELRVGPQDATHCSPGLSCRLWDS DILCLPGDIVPAPGPVLAP
THLQTELVLRQCKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVV
LSFQAYPTARCVLLEVQVPAALVQFGQSVGSSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQ
QLPALPWLNV SADGDNVHLVLNVSEEQHFGLSLYWNQVQGPPKPRWHKNLTG PQIITLNHTDL
VPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLWQAARLRLLTLQSWLLDAPCSLPAEAALCW
RAPGGDPCQPLVPPLSWENVTVDKVLEFP LLKGHPNLCVQVNSSEKLQLQECLWADSLGPLKD
DVLLLETRGPQDNRS LCALEPSGCTSLPSKASTRAARLGEYLLQDLQSGQCLQLWDDDLGALW
ACPMDKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRS GAAARGRAALLLY
SADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAFWFAQRRQTLQEGGVVLLFS
PGAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPA
LFRTVPVFTLPSQLPDFLGALQQPRAPRSGR LQERAEQVSRA LQPALDSYFHPPGTPAPGRGV
GPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394**Glycosaminoglycan attachment site.**

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCG**ATG**AGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGG
 CTGCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCCAGCAACT
 TTGAAAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCCAGACACGGTCTACAGCATCG
 AGTATAAGACGTACGGAGAGAGGGGACTGGGTGGCAAAGAAGGGCTGTAGCGGATCACCCGGA
 AGTCCTGCAACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
 GTCAGTGCGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTTCAGCTCTCTGCAGCACACT
 ACCCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTTCAGATGATTGTTTCAT
 CCTACCCCCACGCCAATCCGTGCAGGCGATGGCCACCGGCTAACCTGGAAGACATCTTCCAT
 GACCTGTTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGAGAGGAAG
 CAGAGAGAATATGAGTTCTTCGGCCTGACCCCTGACACAGAGTTCCTTGGCACCATCATGATT
 TGC GTTCCCACCTGGGCCAAGGAGAGTGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGAC
 CGGACATGGACCTACTCCTTCTCCGGAGCCTTCCTGTTCTCCATGGGCTTCCTCGTCGCAGTA
 CTCTGCTACCTGAGCTACAGATATGTACCAAGCCGCTGCACCTCCCAACTCCCTGAACGTC
 CAGCGAGTCTTGACTTTCCAGCCGCTGCGCTTCATCCAGGAGCACGTCTGATCCCTGTCTTT
 GACCTCAGCGGCCCCAGCAGTCTGGCCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGA
 CCCAGGGAGCCCGCAGGAGCTCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAG
 CCAGACATCTCCATCCTCCAGCCCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCC
 TATGCCCCAAACGCTGCCCCCTGAGGTGCGGCCCCCATCCTATGCACCTCAGGTGACCCCCGAA
 GCTCAATTCCCATTCTACGCCCCACAGGCCATCTCTAAGGTCCAGCCTTCCTCCTATGCCCCCT
 CAAGCCACTCCGGACAGCTGGCCTCCCTCCTATGGGGTATGCATGGAAGGTTCTGGCAAAGAC
 TCCCCCACTGGGACACTTTCTAGTCCTAAACACCTTAGGCCTAAAGGTGAGCTTCAGAAAGAG
 CCACCAGCTGGAAGCTGCATGTTAGGTGGCCTTTCTCTGCAGGAGGTGACCTCCTTGGCTATG
 GAGGAATCCCAAGAAGCAAAATCATTGCACCAGCCCTGGGGATTTGCACAGACAGAACATCT
 GACCCAAATGTGCTACACAGTGGGGAGGAAGGACACCCACAGTACCTAAAGGGCCAGCTCCCC
 CTCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTCCCTCCCTTTGCAACCTCCTTCCGGT
 CCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCCTGCTGGAGTCCCTTGTGTGTCCC
 AAGGATGAAGCCAAGAGCCCAGCCCCCTGAGACCTCAGACCTGGAGCAGCCACAGAACTGGAT
 TCTCTTTTCAGAGGCCTGGCCCTGACTGTGCAGTGGGAGTCT**TGAG**GGGAATGGGAAGGCTT
 GGTGCTTCTCCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCAATCCCATGCCTGCCCAT
 GCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGCAGAGGGAGTGGCATG
 CAGGGCCCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATAAGGACTGCAGCGG
 GGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCGTGTGCTGAGCCCTGCAAGGCAGAAA
 TGACAGTGCAAGGAGGAAATGCAGGGAAACTCCCAGAGTCCAGAGCCCCACCTCCTAACACCA
 TGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGCCCCATTCTGGCCAGTTTCACAATCT
 AGCTCGACAGAGCATGAGGCCCCCTGCCTCTTCTGTCAATTGTTCAAAGGTGGGAAGAGAGCCTG
 GAAAAGAACCAGGCCTGGAAAAGAACCAGAAGGAGGCTGGGCAGAACCAGAACAACCTGCACT
 TCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGCTCAT
 TCCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTTCAGCTTCATTCTCTGATAGAACAAG
 CGAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGAC
 CCTATCCTGAGAATGGGGTTTGAAAGGAAGGTGAGGGCTGTGGCCCCCTGGACGGGTACAATAA
 CACACTGTACTGATGTACAACCTTTGCAAGCTCTGCCTTGGGTTAGCCCATCTGGGCTCAAA
 TTCCAGCCTCACCCTCACAAGCTGTGTGACTTCAAACAATGAAATCAGTGCCAGAACCTC
 GGTTTCTCATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAA
 ATGAAGTCATGTCTTTAAAGTGCTTAATAGTGCCTGGTACATGGGCAGTGCCCAATAAACGGT
 AGCTATTTAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGER
DWVAKKGCQRITRKSCNLTVETGNLTLEYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDV
TCISKVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFF
GLTPDTEFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYR
YVTKPPAPPNSLNVQRVLTFOPLRFIQEHVLI PVFDLSGPSSLAQPVQYSQIRVSGPREPAGA
PQRHSLSEITYLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYA
PQAISKVQPSSYAPQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCM
LGGLSLQEVTSLAMEESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKGQLPLLSSVQI
EGHPMSLPLQPPSGPCSPSDQGPPSPWGLLES LVC PKDEAKSPAPETSDLEQPTELDSLFRGLA
LTVQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCCGCGGTGGCCACAACAT
GGCTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTCGTGCTGGGGGCGCTCTGGTGGGTCCCGGG
CCAGTCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTG
CAGCATGTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGGCCCTGATTGTCGTTTTGT
GAATTTTAAAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAAC
TTGGGCTGGAAGTGTTGAACACAGTTTTGGATATTTTCCAAAAGATTTGATCAAGGTACTTCA
TAAATACACGGAAGAAGAGCTACATATTCCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGG
AGGAAGAGATGATTTTAAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAAGTGGAGGA
CTCTGTACCTGAAGAGTCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGA
GGAGTCTCGGGGGCGTGAACCTTGACCCTGTGCCTGAGCCCGAGGCATTGAGAGCTGATTCAGA
GGATGGAGAAGGTGCTTTCTCAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGA
GAGCCACCCTCACACCAGCGGTCCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACAC
TTTTGAAGAAATTCTGCACGATAAATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAG
TTCTCCTGCCTCGGTGGAGCGGGAGAAGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAG
TCAGAGAGGAAGTGGACAGTGC GTTATTCATTACAGCAAAGGATTTTCGTTGGCATCAAATCT
AAGTTTGT TTTACAAAGATTGTTTTTTAGTACTAAGCTGCCTTGGCAGTTTGCATTTTTGAGCC
AAACAAAAATATATTATTTTCCCTTCTAAGTAAAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFWLGALWWVPGQSDLSHGRRFSDLKVCGDEEC SMLMYRGKALEDFTGPDCRF
 VNFKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLH KYTEELHIPADETDFVCFE
 GGRDDFNSYNVEELLGSLELEDSVPEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRADS
 EDGEGAFSESTEG LQGQPSAQESH PHTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGN
 SSPASVEREKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTGAGAGGCCGGGGAAGAGAAGCAAAGC
GCAACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCCCTAA
CTTCAGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGCGG
CACAGGCGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTTCGGGCAGCTGG
GCTCGGGCGGCGGGAGTAGGGCCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGG
CTGCGCCCTGGGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGA
TGAGCCGCGTGGTCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCC
GCCGCGTGGTCAAGCGGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGG
CCTACTTCCATGAACTGTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTG
AGGGAGGAGTCCTCCTCAGCCTTGAGAATGAAGCAGAACAGAAAGTTAATAGAGAGCATGTTGC
AAAACCTGACAAAACCCGGGACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGA
ATGGAGATGGGCAAACATCTGGTGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATT
CCCAGTACCGAAACTGGTACACAGATGAACCTTCCTGCGGAAGTGAAAAGTGTGTTGTGATGT
ATCACCAACCAACTGCCAATCCTGGCCTTGGGGGTCCCTACCTTTACCAGTGGAAATGATGACA
GGTGTAACATGAAGCACAATTATATTTGCAAGTATGAACCAGAGATTAATCCAACAGCCCCCTG
TAGAAAAGCCTTATCTTACAAATCAACCAGGAGACACCCATCAGAATGTGGTTGTTACTGAAG
CAGGTATAATTCCCAATCTAATTTATGTTGTTATACCAACAATACCCCTGCTCTTACTGATAC
TGGTTGCTTTTGGAACTGTTGTTTCCAGATGCTGCATAAAAGTAAAGGAAGAACAAAACTA
GTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAAGAAAGTGGCATGGAAGTAT
AATAACTCATTGACTTGGTTCCAGAATTTTGTAACTCTGGATCTGTATAAGGAATGGCATCAG
AACAATAGCTTGGAAATGGCTTGAAATCACAAAGGATCTGCAAGATGAACTGTAAGCTCCCCCT
TGAGGCAAATATTAAAGTAATTTTATATGTCTATTATTTTCAATTTAAAGAATATGCTGTGCTA
ATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAACTTCAAACCTTCAAGCAAA
TGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTATGTGTGTTAGAAGCAAT
TCCTTTTATTTCTTTTACCTTTTACATAAGTTGTTATCTAGTCAATGTAATGTATATTGTATTGA
AATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGATAAAAATGAACTGTTCTA
ATATTTATTTTATGGCATCTCATTTTCAATACATGCTCTTTTGATTAAAGAACTTATTAC
TGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGTCTCTCGAA
ATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGAAATAAGA
AGCTATTTTCAATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGATTGTCT
AATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTACTTAAATTAACAGATTTTG
TAATAATGTAACCTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAGT
GACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTG
AGGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTTCGGGGTTT
GGGATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGC
CTCTGACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAG
TGTGGAGACAAGCACAGCACACAGACATTTTAGGAAGGAAGGAACACTACGAAATCGTGTGAAA
ATGGGTTGGAACCCATCAGTGATCGCATATTCATTGATGAGGGTTTGCTTGAGATAGAAAATG
GTGGCTCCTTTCTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCTTGTTCTTCTCAAGAGA
AAGTTGTAACCTCTCTGGTCTTCATATGTCCCTGTGCTCCTTTTAAACCAATAAAGAGTTCTTG
TTTCTGGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFRRVVSQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACES
EGGVLLSLENAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSN
SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNKYCKYEPEINPTAP
VEKPYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKT
SPNQSTLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217